Hepatocyte Nuclear Factor 4-Alpha Is Essential for the Active Epigenetic State at Enhancers in Mouse Liver

Thakur, A., Wong, J. C. H., Wang, E. Y., Lotto, J., Kim, D., Cheng, J. C., Mingay, M., Cullum, R., Moudgil, V., Ahmed, N., Tsai, S. H., Wei, W., Walsh, C. P., Stephan, T., Bilenky, M., Fuglerud, B. M., Karimi, M. M., Gonzalez, F. J., Hirst, M., & Hoodless, P. A. (2019). Hepatocyte Nuclear Factor 4-Alpha Is Essential for the Active Epigenetic State at Enhancers in Mouse Liver. *Hepatology*, *70*(4), 1360-1376. https://doi.org/10.1002/hep.30631

Published in:
Hepatology

Publication Status:
Published (in print/issue): 08/10/2019

DOI:
10.1002/hep.30631

Document Version
Author Accepted version

HNF4A is essential for the active epigenetic state at enhancers in mouse liver

Avinash Thakur[1,3], Jasper C. H. Wong[2], Evan Y. Wang[1], Jeremy Lotto[1], Donghwan Kim[4], Jung-Chien Cheng[1], Matthew Mingay[2], Rebecca Cullum[1], Vaishali Moudgil[1], Nafeel Ahmed[1], Shu-Huei Tsai[1], Wei Wei[1], Colum P. Walsh[5], Tabea Stephan[1], Misha Bilenky[6], Bettina M. Fuglerud[1,7], Mohammad M. Karimi[8], Frank J. Gonzalez[4], Martin Hirst[2,6] and Pamela A. Hoodless[1,3,9]

[1]Terry Fox Laboratory, BC Cancer, Vancouver, British Columbia, Canada, V5Z 1L3

[2]Department of Microbiology and Immunology, Michael Smith Laboratories Centre for High-Throughput Biology, University of British Columbia, Vancouver, British Columbia, Canada

[3]Department of Medical Genetics, University of British Columbia, Vancouver, Canada, V6T 1Z4

[4]Center of Cancer Research, National Cancer Institute, Bethesda MD 2089

[5]Genomic Medicine Research Group, Centre for Molecular Biosciences, Biomedical Sciences Research Institute, Ulster University, Coleraine, BT52 1SA, UK.

[6]Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada

[7]Department of Biosciences, University of Oslo, Oslo, Norway, 0316

[8]MRC Institute of Medical Sciences, Imperial College London, W12 0NN, UK

[9]School of Biomedical Engineering, University of British Columbia, Vancouver, Canada, V6T 1Z4

Keywords: Histone modification, hydroxymethylation, hepatoblasts, hepatocytes, TET3

Contact address:

Dr Pamela A. Hoodless

Terry Fox Laboratory, BC Cancer

675 West 10th Avenue, Vancouver, British Columbia

Canada, V5Z 1L3

Fax: 604-877-0712

Email: hoodless@bccrc.ca

List of Abbreviations:

HNF4A: Hepatocyte nuclear factor 4-alpha

TF: Transcription Factor

FOXA2: Forkhead Box A2

TET3: Ten-eleven translocation methylcytosine dioxygenase 3

HBR: HNF4A bound region

FBR: FOXA2 bound region

5mC: 5-methylcytosine;

5hmC: 5-hydroxymethylcytosine

PLA: Proximity ligation assay

Financial support: This work was funded by Genome BC and the Canadian Institutes of Health Research (FRN128092) to P.A.H and F.J.G is supported by National Cancer Institute Intramural Research Program.

ABSTRACT

Cell fate determination is influenced by interactions between master transcription factors (TFs) and cis-regulatory elements. Hepatocyte nuclear factor 4 alpha (HNF4A), a liver-enriched TF, acts as a master controller in the specification of hepatic progenitor cells by regulating a network of TFs that controls the onset of hepatocyte cell fate. Using analysis of genome-wide histone modifications, DNA methylation and hydroxymethylation in mouse hepatocytes, we show that HNF4A occupies active enhancers in hepatocytes and essential for active histone (H3K27ac) and DNA signatures, especially 5-hydroxymethylcytosine (5hmC). In mice lacking HNF4A protein in hepatocytes, we observed a decrease in both H3K27ac and hydroxymethylation at regions bound by HNF4A. Mechanistically, HNF4A-associated hydroxymethylation (5hmC) requires its interaction with TET3, a protein responsible for the oxidation from 5mC to 5hmC. Furthermore, HNF4A regulates TET3 expression in liver by directly binding to an enhancer region. In conclusion, we identified that HNF4A is required for the active epigenetic state at enhancers that facilitates transcription from genes in hepatocytes.

INTRODUCTION:

Hepatocyte nuclear factor 4A (HNF4A, also known as NR2A1) is a member of the nuclear receptor superfamily of transcription factors (TFs) and is expressed in the liver, pancreas, gastrointestinal tract and kidneys [1]. In the liver, HNF4A is required for normal liver development [2], including morphogenesis [3] and differentiation [4], and controls a set of genes necessary for hepatocyte functions. HNF4A deficiency in the mouse embryonic liver causes failure in hepatocyte differentiation [2], while in adult mice loss of HNF4A affects lipid homeostasis resulting in liver steatosis [5]. In humans, HNF4A mutations have been linked to maturity-onset diabetes of young type1 (MODY-1), a monogenic form of type 2 diabetes [6]. In addition to maintaining the expression of various genes associated with metabolic processes in the liver, several studies have connected reduced HNF4A expression with liver cancer progression [7]. HNF4A disruption in mature hepatocytes has been shown to result epithelial to mesenchymal transition (EMT), a process associated with cancer progression [8]. Conversely, its overexpression in hepatocellular carcinoma rodent models shows its role in tumour suppression [4, 7] by repressing the proliferative capacity of cancer cells[9]. Taken together, these studies strongly support a role for HNF4A as a central regulator of hepatocyte identity by playing a crucial function in regulating liver gene expression [5, 10]. However, the mechanisms through which HNF4A exerts its functions are not well understood.

TFs, such as HNF4A, primarily function through directly binding to DNA at sequence-specific motifs; however, the number of potential binding sites for a given TF across the genome far exceeds the binding events observed *in vivo*. We have shown that HNF4A binding is highly dynamic and that binding events can change based on tissue and the differentiation state [11]. Genome-wide analysis of mammalian genomes has provided evidence that active cis-regulatory regions are located in open chromatin regions [12], while condensed chromatin is inaccessible

and thus inactive. How chromatin structure is regulated in tissue- and differentiation stage-specific patterns and how TFs play a role in the adaptation of the chromatin structure is only beginning to be explored.

DNA methylation and histone modifications are two of the primary mechanisms of epigenetic control that influence chromatin condensation and DNA accessibility. Histone modifications that include methylation and acetylation of histone tails have been associated with repressed, poised and active enhancers and promoters, and ultimately control gene activity [13, 14]. In particular, lysine 4 of histone 3 monomethylation (H3K4me1) and trimethylation (H3K4me3) mark enhancers and promoters respectively [15]. Acetylation of lysine 27 of histone 3 (H3K27ac) is also associated with the active state of the element.

In addition to histone modifications, DNA methylation (5-methylcytosine or 5mC) is associated with gene repression [16] and can disrupt the binding of many TFs [17, 18]. DNA hydroxymethylation or 5-hydroxymethylcytosine (5hmC) is a recently discovered DNA modification that is produced by the oxidation of 5mC due to the activity of ten-eleven translocation proteins (TET1, TET2 and TET3). 5hmC plays an important role during mammalian development and cellular differentiation since *Tet1/2/3* deficiency in ES cells (triple knockout cells) reduces their differentiation capabilities and also limits their developmental potential during embryogenesis [19]. Although 5hmC was shown to be highly enriched at gene bodies in brain tissue, genome-wide enrichment studies in human and mouse embryonic stem cells (ESCs) confirms the presence of 5hmC at promoters and enhancers in addition to gene bodies [20].

A mechanistic understanding of the interrelationship between TF binding, histone modification and DNA methylation states is only beginning to emerge. The establishment and maintenance of tissue-specific regions of histone and DNA modifications may be controlled by

TFs that can directly or indirectly interact with epigenetic modifiers [21], such as the histone acetylase CBP/p300 [22]. Recent reports have identified TFs that interact with TET proteins to maintain low 5mC and high 5hmC at enhancers [23, 24].

Studies transducing a combination of TFs into human fibroblasts to produce highly functional hepatocytes support HNF4A's involvement in chromatin remodelling. These studies revealed that HNF4A is crucial for generating induced hepatocyte-like cells (iHeps) [25, 26] as HNF4A overexpression alone may be sufficient to generate these cells [27]. These experiments suggest that HNF4A may help establish an appropriate epigenetic state to allow binding of additional TFs that regulate hepatocyte-specific gene expression. Here, we show that HNF4A maintains an active epigenetic state at cis-regulatory elements in hepatocytes and during liver differentiation. Loss of HNF4A during liver differentiation confirmed that HNF4A is required for establishment and maintenance of H3K27ac. Furthermore, HNF4A-mediated oxidation of 5mC during differentiation of hepatoblasts to hepatocytes is achieved by its interaction with TET3, a newly identified HNF4A target gene in hepatoblasts. Overall, we confirm that HNF4A-driven epigenetic changes are required for expression from its target genes during hepatocyte differentiation. These findings provide a novel mechanism through which HNF4A works as a master regulator to shape and maintain the epigenetic landscape in the liver.

MATERIALS AND METHODS

Mice and cell preparation

Mouse protocols were approved by the University of British Columbia Animal Care committee. Livers were collected from 8 week old female C57BL6/J mice, perfused with 0.005% collagenase (STEMCELL Technologies) in PBS, minced with a razor blade and passed through a 40μm cell strainer. Hepatocytes were separated by low-speed centrifugation (50g x 30sec) [28].

Hepatoblast isolation from E14.5 livers was previously described [11]. *Alb*-HNF4$^{F/F;AlbERT2cre}$ and *Hnf4a*$^{F/F}$ mouse were described in an earlier study [4]. These mice were crossed with Tamoxifen inducible cre recombinase (specific to hepatocytes) expressing mouse. Mice used in this study were 6-8 weeks old and were treated with Tamoxifen (2mg/ml, intraperitoneal, subcutaneously) every other day (total 3 injections) for 7 days. Mice were killed by cervical dislocations under anesthesia and livers were collected 7th day post-injection. For HNF4$^{F/F;AlbERT2cre}$ mice, floxed cre negative mice without Tamoxifen were used as control. The HNF4a knockout animal studies and procedures were carried out in accordance with the Institute of Laboratory Animal Resources guidelines and approved by the National Cancer Institute Animal Care and Use Committee. All mice were housed in a pathogen-free animal facility under a standard 12-hour light/dark cycle and given pelleted NIH-31 chow diet (at NIH) or Envigo-Teklad 2920X (at BC Cancer) and water *ad libitum*.

5mC and 5hmC Analysis

Low input MeDIP-seq and hMeDIP-seq was previously described [29]. DNA was sonicated to 200-400bp using E-220 sonicator (Covaris) followed by end repair adapter ligation, and immunoprecipitation using anti-5mC or anti-5hmC antibodies. Two PCR reactions of 12 cycles using paired-end, indexed, Illumina PCR primers amplified the DNA for library preparation. Size selection of fragments <700bp was achieved on 8% polyacrylamide gels. Paired-end, 125 nucleotide reads were generated on Illumina MiSeq or Illumina Hiseq 2500 platform (v3 chemistry). Bioinformatic analysis of MeDIP and hMeDIP-seq data was performed as described in the supplementary material and methods section. MeDIP and hMeDIP for locus-specific qPCR analysis were performed as described in Supplementary materials and methods.

Histone modification and transcription factor analysis

ChIP was performed as previously described [30] with the following modifications. Cells or minced tissues were fixed with 1% formaldehyde (Sigma) for 10 minutes then quenched using 125 mM glycine for 5 minutes at room temperature. MACs purified DLK1+ cells were fixed prior to cell sorting. Cells were lysed for 15 min on ice in lysis buffer (10mM Tris-HCl pH8, 10mM NaCl, 3mM MgCl2, 0.5% Nonidet P-40). Nuclei were pelleted and lysed for 30-60 min on ice with nuclear lysis buffer (50 mM Tris-HCl pH 8, 5 mM EDTA, 1% SDS). Samples were sonicated with Q Sonica Sonicator Q700 for 10 min total in 30 second cycles to generate fragments of 200-500bp. Chromatin was diluted in ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA) and Protein A/G beads (ThermoFisher) was added to clear the samples prior to addition of antibody. Antibody (3µg for transcription factor and 2ug for histone modifications) was added to samples and incubated overnight at 4°C. Protein A/G beads (ThermoFisher) were added and samples were placed on a rotator at RT for 4 hours. The samples were washed with low salt (20 mM Tris-HCl pH 8, 0.15 M NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), high salt (20 mM Tris-HCl pH 8, 0.5 M NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), lithium chloride (10 mM Tris-HCl pH 8, 0.25 M LiCl, 1 mM EDTA, 1% Nonidet P-40, 1% sodium deoxycholate) and TE buffer for 10 minutes each followed by elution with 0.1M sodium hydroxycholate in 1% SDS. DNA purification and qPCR analysis were the same as above. For ChIP-seq library construction and Illumina Genome Analyzer sequencing was performed as described [15]. Bioinformatic analysis of ChIP-seq data was performed as described in the Supplementary material and methods section.

Cell culture and reagents

144 HepG2 and HEK293T cells were purchased from American Type Culture Collection (Manassas,
145 VA). Cells were grown in DMEM media (STEMCELL Technologies) supplemented with 10%
146 fetal bovine serum (HyClone Laboratories, Logan, UT). HPPL cells were generated and
147 maintained as previously described [31]. All cells were maintained in a humidified incubator at
148 37°C and 5% CO_2. For knockdowns, cells were seeded on 6 well plates and transfected with
149 HNF4A and TET3 siRNA (Dharmacon) using RNAiMAX (ThermoFisher) for 72 hrs. siRNA used
150 in this study are presented in Table S2.
151
152 **Co-immunoprecipitation (co-IP)**
153 HEK293T cells were transfected with an HNF4A expression vector [32] and HA-tagged TET3
154 vector (Addgene #49446) alone and in combination using Polyethylenimine (PEI). Cells were
155 lysed using lysis buffer (10mM Tris-HCl pH8, 10mM NaCl, 3mM MgCl2, 0.5% Nonidet P-40) and
156 nuclei were collected on ice. CHAPS buffer was used for nuclear extraction and the lysate was
157 incubated with 2ug of HNF4A antibody overnight at 4°C on a rocker. Protein A/G beads
158 (ThermoFisher) were added (20ul) and samples were rotated for 30 minutes at RT. Beads were
159 collected by centrifugation and extensively washed with lysis buffer. Proteins were eluted into
160 SDS gel loading buffer by heating at 65°C. Proteins were separated using NuPAGE 4-12% gels
161 (ThermoFisher), transferred to PVDF membranes, and analysed using the antibodies specified
162 in Table S2.
163
164 ***In situ* proximity ligation assay (PLA)**
165 Embryonic (E14.5) and adult (8 weeks old) mouse livers were fixed with 4% paraformaldehyde
166 (Sigma) at 4°C for 16-48 hrs, put through a sucrose gradient at 4°C, and embedded in OCT
167 (TissueTek). Sections (4 μm) were permeabilized with Triton-X (0.1% v/v in PBS). PLA was

168 performed according to manufacturer's instructions using the Duolink *In situ* PLA Kit (Sigma
169 Aldrich). Fluorescent foci were imaged using the Zeiss Axio Imager M2. Cells were seeded on
170 coverslips and washed twice with PBS before fixation and permeabilization.
171
172 **Statistical analysis**
173 Results are calculated as mean ±SEM derived from at least three independent experiments.
174 Student's *t-test* was used for qPCR data with GraphPad software. A p-value <0.05 was
175 considered statistically significant.
176
177 **RESULTS**
178 **Regions bound by HNF4A are associated with active histone and DNA modifications in**
179 **hepatocytes**
180 HNF4A is a well-studied master regulator of hepatocyte identity; however, how it contributes to
181 epigenetic plasticity in the liver is not very well understood. To address this, we investigated
182 epigenetic modifications associated with HNF4A binding in adult mouse hepatocytes. We used
183 our previously published HNF4A [11], H3K4me1 and H3K4me3 [15], and newly generated
184 H3K27ac ChIP-seq datasets and examined the enrichment of histone modifications around
185 HNF4A bound regions (HBRs) (Fig. 1A). HNF4A antibody used for ChIP-seq (Table. S2)
186 identifies both embryonic and adult isoforms. Analysis of HNF4A ChIP-seq data revealed its
187 preferential binding to intronic (47.35%), and distal intergenic regions (33.33%) with a small
188 proportion at gene promoters (9.81%) (Fig. S1A). Examination of HBRs demonstrated a high
189 enrichment of H3K4me1 and H3K27ac flanking these regions (Fig. 1A), suggestive of active
190 enhancers. The active status of these enhancers was further confirmed by examination of
191 mouse liver DNase-seq data (ENCODE), which shows very high signals, demonstrative of open

chromatin at HBRs (Fig. S1B). It is well established that active enhancers are associated with high 5hmC and low 5mC DNA modifications [14, 33]. To examine the level of 5mC and 5hmC at HBRs, we isolated hepatocytes from mouse liver and performed MeDIP (Methylated DNA Immunoprecipitation) and hMeDIP (5-hydroxymethylated DNA Immunoprecipitation) followed by high-throughput sequencing. We observed that HBRs are enriched for 5hmC but depleted of 5mC (Fig. 1A) as illustrated by the combined profile plot (Fig. 1B). The epigenetic status of HBRs, such as those near the genes *Cdc42bpb*, *Rfx4* (Fig. 1C and D), *Mgst3* and *Ido2* (Fig. S1C & D), demonstrate the features of active enhancers. To validate our sequencing data, we confirmed 5mC depletion and 5hmC enrichment by MeDIP- and hMeDIP-qPCR at *Cdc42bpb*, *Rfx4, Mgst3, Ido2*, and *Tet3* HBRs (Fig. 1E). *Dazl*, *Tbx15* and *IGd* were included as positive, negative and no CpG controls, respectively. As expected, the level of 5mC at CpG islands associated with *Dazl* and *Tbx15* was very similar to what we observed in our sequencing data (Fig. S1E and F). Overall, our data show that HBRs are enriched for H3K4me1 and H3K27ac, DNaseI hyper-sensitivity and 5hmC, which together indicate that HNF4A is primarily bound at active enhancers in hepatocytes.

Regions bound by HNF4A do not require FOXA2 to have a transcriptionally permissive epigenetic state in hepatocytes

We previously reported a significant overlap between HNF4A and FOXA2 bound regions (FBRs) in hepatocytes [11]. As a pioneer factor, FOXA family members are capable of altering epigenetic modifications to regulate gene expression [24, 34]. We, therefore, speculated that a high level of H3K4me1 and H3K27ac at HBRs could be due to FOXA2 occupancy at those sites. To address this, we performed K-means clustering (K=5) on HNF4A and FOXA2 bound regions in hepatocytes using ChAsE [35] and identified 5 clusters. As expected, a large overlap

was observed between HBRs and FBRs (Cl2, Cl3 and Cl4), however, two clusters showed exclusive binding of either HNF4A or FOXA2 (Cl1 and Cl5, respectively) (Fig. 2A). Further, these clusters revealed distinctive histone and DNA signature patterns with Cl1 being enriched for H3K4me1, H3K27ac and 5hmC but depleted for 5mC compared to Cl5 (Fig. 2A). Given the importance of CpG density and its impact on 5mC, we performed CpG density analysis and found no significant difference between clusters, indicating 5mC enrichment in Cl5 is independent of CpG density (data not shown). To get a deeper insight into the difference in histone and DNA modifications associated with these clusters, we generated profile plots for H3K4me1 (Fig. 2B), H3K27ac (Fig. 2C), 5hmC (Fig. 2D), and 5mC (Fig. 2E). Profile plots identify the decreased level of H3K4me1, H3K27ac and 5hmC but an increase in 5mC in the cluster lacking HNF4A (Cl5). Furthermore, Cl5 regions show the least chromatin accessibility of all clusters based on DNaseI-seq data from mouse liver (Fig. S2A). Cluster 2, 3 and 4 displayed almost identical patterns of histone and DNA modifications. To explore whether there is a functional consequence of these differing chromatin states across clusters, we looked at gene expression for the single nearest gene transcriptional start site (TSS) within 20kb using RNA-seq data from hepatocytes [11]. Interestingly, genes associated with Cl5 have a significantly lower expression compared to all other clusters (Fig. 2F). The other four clusters, which all have bound HNF4A, had similar high expression levels characteristic of active transcriptional elements. Analysis of Cl1 confirmed no enrichment of motifs for FOXA in those regions suggesting no contribution of FOXA family members in maintaining a transcriptional permissive state at those regions (Fig. S2B). In contrast, motif analysis of cluster 5 revealed high enrichment for FOXA2 and low levels of enrichment for NFIC, HNF4A and CTCF (Fig. S2C). Interestingly, we observed low enrichment of HNF4A motifs in cluster 5 but our ChIP-seq data on HNF4A confirmed no binding at those locations. Pathway analysis to identify the functions of

genes associated with these clusters showed that the cluster 1 genes are primarily associated with metabolic functions (Fig. S2D) and cluster 5 genes are associated with various signalling pathways (Fig. S2E). Overall these findings, confirms that regions bound by HNF4A, irrespective of presence or absence of FOXA2, are accessible and actively enhancing transcription of genes primarily associated with metabolic functions in liver.

Without HNF4A, H3K27ac and 5hmC are significantly decreased at HBRs in hepatocytes

With the observation that H3K27ac and 5hmC are enriched at HNF4A bound regions, we hypothesized that HNF4A is required for appropriate enhancer modifications that ensure precise expression requirements independent of FOXA2. To test this, we took advantage of a Tamoxifen-inducible, hepatocyte-specific, HNF4A knockout (KO) mouse [4] and performed ChIP-, MeDIP- and hMeDIP-qPCRs on WT and KO livers. For our analysis, we investigated sites bound by both HNF4A and FOXA2 (Cl2, Cl3 and Cl4, see Fig. 3A for example) as well as sites bound by HNF4A alone (Cl1, see Fig. 3B for example) for changes in H3K27ac patterns. H3K27ac ChIP-qPCR in WT and KO livers showed that loss of HNF4A resulted in decreased H3K27ac at all sites, indicating a requirement for HNF4A in ensuring H3K27ac is maintained. Positive and negative controls that are not normally bound by HNF4A in the liver (*Gapdh* and *Six1*) had no significant changes in H3K27ac levels (Fig. 3C). Furthermore, in KO livers, the same genomic locations exhibited a reduction in 5hmC to same extent (almost two fold change) for both HNF4A only and HNF4A/FOXA2 shared regions (Fig. 3D). We also observed a small increase in 5mC levels in HNFKO livers at most loci examined (Fig. S3A). This investigation was extended *in vitro* using siRNA-mediated knockdown in HepG2 cells. We confirmed a reduction of HNF4A mRNA and protein expression (Fig. S3B) and validated that HNF4A bound to regions in HepG2 cells as identified in ENCODE data (e.g., *RBKS*, *MGST2*, *SLCO4C1*,

APOA4, *PRLR* and *AGXT*) (Fig. S3C). After siRNA depletion of HNF4A, these sites had significantly decreased levels of 5hmC (Fig. S3D) and increased levels of 5mC (Fig. S3E). Together, these findings suggest that HNF4A is required to maintain high levels of H3K27ac and 5hmC in cell lines and *in vivo*.

The gain of HNF4A during differentiation results in increased H3K27ac and 5hmC at HBR associated enhancers

We have previously reported the phenomenon of enhancer switching during liver differentiation where TFs, including HNF4A, show distinct binding sites in embryonic hepatoblasts and adult hepatocytes [11]. To determine if H3K27ac is altered in response to differential binding of HNF4A in hepatoblasts and hepatocytes, we compared ChIP-seq data for HNF4A and H3K27ac in hepatoblasts purified from E14.5 embryonic livers and in adult hepatocytes. We identified 2765 hepatocyte-enriched HBRs that are present in differentiated hepatocytes but not bound in hepatoblasts (examples are shown in Fig. 4A left and right). Of note, these hepatocyte-enriched HNF4A binding sites showed a gain of H3K27ac (Fig. 4B) in hepatocytes compared to hepatoblasts. In contrast, sites at which HNF4A-binding is lost after differentiation (hepatoblast-enriched) showed decreased H3K27ac in adult hepatocytes (Fig. S4A). Moreover, hepatocyte-enriched HBRs show decreased levels of 5mC (Fig. 4C) and increased levels of 5hmC (Fig. 4D) in hepatocytes compared to hepatoblasts, suggesting that the gain of HNF4A in those regions results in the oxidation of 5mC to 5hmC. Notably, genes associated with HNF4A sites gained in hepatocytes were significantly more highly expressed in hepatocytes compared to hepatoblasts (individual and global examples are shown in Fig. 4E and 4F), whereas genes associated with regions no longer bound by HNF4A in the hepatocytes showed the opposite trend (Fig. S4B).

Overall these findings support an important role of HNF4A in cell type-specific gene regulation by ensuring high levels of H3K27ac and 5hmC at gene regulatory regions.

HNF4A and TET3 proteins physically interact

We speculated that HNF4A interacts with TET proteins to maintain low-levels of 5mC at bound regions through oxidation to 5hmC. We focused on TET3 due to its high expression in hepatocytes (Fig. 5A and 5B) and its known interactions with another nuclear receptor family member, TRα1 [36]. Of note, TET3 co-immunoprecipitation with HNF4A in HepG2 cells, which express high levels of both HNF4A and TET3, confirmed that the two proteins interact (Fig. 5C). This interaction was also confirmed using the expression of tagged versions of the proteins in HEK293T cells (Fig.SB). Importantly, the TET3-HNF4A interaction was observed in nuclear lysates isolated from mouse hepatocytes (Fig. 5D) as well as hepatoblasts (Fig. 5E). Proximity ligation assay (PLA) on adult and embryonic mouse liver cryosections validated the close association between HNF4A and TET3 (Fig. 5F) in nuclei of hepatoblasts and hepatocytes *in vivo*. These PLA findings were further validated in HEK293T cells using the expression of Myc-HNF4A and Flag-TET3 (Fig. S5D). Together, our findings suggest that HNF4A physically interacts with TET3, *in vivo,* to mediate the active 5hmC status at regulatory regions to facilitate gene transcription. These findings suggest that the presence of both TET3 with HNF4A is required for expression of genes associated with HBRs.

TET3 is a direct transcriptional target of HNF4A

In hepatoblasts, we observed a strong intragenic HNF4A binding site approximately 19kb downstream from the *Tet3* transcriptional start site. This HNF4A site co-localizes with active histone modifications (H3K4me1 and H3K27ac), suggesting an active enhancer (Fig. 6A).

Knockdown of HNF4A in mouse HPPL cells (hepatic progenitor cells proliferating on laminin), which are capable of generating hepatocyte and cholangiocytes similar to mouse hepatoblasts, caused a reduction of *Tet3* mRNA expression ($p<0.0001$) and protein levels (Fig. 6B), suggesting that HNF4A regulates *Tet3* expression in hepatic cells. To determine if the HNF4A bound enhancer in *Tet3* (mTET3E, a 405bp region) is functional; we used luciferase reporter assays in HEK293T cells. HNF4A overexpression significantly increased ($p>0.0001$) enhancer activity of mTET3E (Fig. 6C). We confirmed the luciferase results using HepG2 cells that endogenously express HNF4A (Fig. 6D). Furthermore, ENCODE ChIP-seq data in HepG2 cells demonstrated a similar HNF4A bound region with active histone modifications between exon 2 and 3 of *TET3* in the human genome (Fig. S6B). Luciferase assays with the human region (332bp) in HEK293T cells with HNF4A overexpression (Fig. S6C), and in HepG2 cells with endogenous HNF4A (Fig. S6D) confirmed HNF4A-mediated regulation of *TET3*.

The mouse TET3 enhancer contains two HNF4A motifs (HNF4A-1 and HNF4A-2) (Fig. 6E). To test which of the HNF4A motif conveyed the enhancer activity, the motifs were deleted sequentially in the reporter construct containing the mouse enhancer and the activity was measured in both HepG2 (Fig. 6F) and HEK293T cells (Fig. S6A). As expected, combinatorial deletion of these motifs diminished luciferase activity in HepG2 cells (Fig. 6F). Of note, in HEK293T cells, the deletion of the two motifs did not have any additional effect compared to the single deletions, suggesting that the hepatic environment of the HepG2 cells may contain additional factors required for this regulation that are absent in the HEK293T cells. Overall, our results suggest that HNF4A directly regulates *Tet3* in the hepatic cells and TET3 then, in turn, interacts with HNF4A to control the epigenetic environment of regions by ensuring 5hmC and H3K27ac are maintained.

DISCUSSION

HNF4A is a key TF in the liver as its loss is associated with transcriptional repression of genes [5, 10]. Although various TFs were shown to regulate gene expression by influencing histone [24, 37] and DNA modifications at gene regulatory regions [24, 37, 38], the role of HNF4A in epigenetic-mediated gene regulation in liver is not well defined.

Here, we show that HNF4A, a master regulator of hepatocyte differentiation, controls gene expression through its influence on epigenetic modifications in hepatoblasts and hepatocytes. In particular, we show that HNF4A binding in hepatocytes is almost exclusively associated with active histone modifications (H3K4me1 and H3K27ac) and DNA hydroxymethylation (5hmC) at enhancers, supporting a primary role of HNF4A in transcriptional activation in the liver. Moreover, our data suggest that HNF4A is required for the presence of H3K27ac and 5hmC since the loss of HNF4A in adult liver, through an inducible knockout model, reduces these active enhancer signatures at HNF4A bound regions. In addition, regions bound by HNF4A in hepatoblasts but not hepatocytes have reduced levels of H3K27ac and 5hmC in hepatocytes. This loss correlates with reduced transcription from the associated genes (Fig. 7). Finally, HNF4A bound regions acquired across differentiation in hepatocytes are enriched for H3K27ac and 5hmC in hepatocytes. Together this data establishes HNF4A as a facilitator of the active enhancer state.

Previous studies have suggested that HNF4A plays multiple roles in chromatin reorganization: ectopic HNF4A expression in fibroblasts was shown to induce DNase1 hypersensitivity of the *Serpin* gene cluster at 14q32.1 to transcriptionally regulate *α1-antitrypsin* and *corticosteroid binding globulin* [39]. This supports a role of HNF4A in the establishment of active enhancers[39]. In addition, HNF4A may regulate chromatin indirectly by controlling the

358 expression of chromatin remodelling genes, such as *Smarcd3* and *Cdt1,* as their expression is
359 altered in HNF4A KO livers [40].

360 The mechanisms through which HNF4A alters the epigenome are likely to be complex.
361 HNF4A was shown to directly interact with CBP/p300, which may account for its ability to
362 regulate H3K27ac [2, 41, 42]. Our data suggest that in the absence of HNF4A, histone acetylases
363 either are not recruited or not able to function at HNF4A enhancers, despite the potential
364 presence of other TF binding. For example, HNF4A frequently co-binds to regions with FOXA2;
365 however, at regions that are bound by FOXA2 alone very little H3K27ac is evident. In addition,
366 in HNF4A KO livers, H3K27ac was reduced at regions bound by either HNF4A only or
367 FOXA2/HNF4A, indicating HNF4A is primarily responsible for maintaining active epigenetic
368 signatures at these locations. In addition to FOXA family members, HNF4A was previously
369 shown to recruit several coactivators such as glucocorticoid receptor interacting protein 1,
370 steroid receptor coactivator1 and cAMP response element binding proteins to regulate gene
371 expression [41, 43]. The role of these co-factors is not clear. Moreover, we observed that RXRA,
372 ESRB and HINFP motifs were enriched in the HNF4A bound regions which lack FOXA2. Thus,
373 H3K27ac at HNF4A binding sites could be affected by interactions with other TFs.

374 In addition to H3K27ac, we observed depletion of 5mC and enrichment of 5hmC around
375 HNF4 bound regions. In HNF4A KO livers, 5hmC was reduced at HBRs. A recent study
376 investigating the role of HNF4A depletion on DNA modifications showed that HNF4A loss can
377 impact global 5mC and 5hmC to a small extent [44]. By first identifying HNF4A-bound regions,
378 we were able to observe significant changes in 5hmC upon HNF4A depletion both *in vivo* and *in*
379 *vitro*, demonstrating its role in the maintenance of this active mark at enhancers (Fig. 7).
380 Interestingly, we observed only a small increase in 5mC at HBRs in HNF4A KO livers,
381 suggesting that either recovery of 5mC is slower than 5hmC loss or it does not occur. A

previous study observed a similar effect at FOXA1 bound enhancers in LNCaP prostate cancer cells in which knockdown of FOXA1 reduced 5hmC without gain of 5mC [24]. Furthermore, it has previously been shown that 5hmC abundance represents only ~10% of the 5mC level in embryonic stem cells,[45], thus changes in 5mC upon HNF4A depletion in the whole liver may be difficult to observe. However, we did observe a clear gain in 5mC in HepG2 cells upon HNF4A knockdown (Fig. S3E).

In the present study, we show that HNF4A functions to regulate DNA methylation in the surrounding genomic regions through interactions with TET3. Of interest, the HNF4A binding motif does not contain a CG which can be methylated, thus making it a methylation-insensitive TF [11, 17]. Thus, HNF4A may bind to its motif and then recruit TET proteins to oxidize the 5mC to 5hmC in the surrounding regions. TET2 and TET3 are known to be required for this 5mC oxidation process in hepatocytes [46]. Here, we focused on TET3 because it is more highly expressed compared to other TETs in the liver and is known to interact with a TF, the nuclear receptor protein TRα1[36]. Since TET2 is also expressed in liver, HNF4A may also interact with TET2 *in vivo*. It should be mentioned that while we focused on 5hmC at HBRs, other DNA modifications, such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) may also be present and important. A recent study showed that TET3 can identify and facilitate removal of 5caC by base excision repair to achieve demethylation of gene regulatory regions [47]. Thus, future work is needed to examine the entire 5mC oxidation pathway in the liver.

In summary, we show that HBRs are associated with a high level of H3K27ac and 5hmC, likely due to HNF4A interactions with CBP/p300 and TET3 respectively (Fig. 7). We identified clusters (Fig. 2A and 7) where loss of HNF4A affects H3K27ac and 5hmC levels to different extents (Fig. 3 and 7). We also confirmed that the loss of HNF4A in a human hepatoma cell line, HepG2, produced similar changes in DNA modifications as observed in *Hnf4a* KO

mouse livers, indicating similar mechanisms of gene regulation (Fig. S3 and Fig. 7).Thus, we have identified a mechanism through which HNF4A regulates epigenetic plasticity in hepatocytes. We have also confirmed HNF4A function in establishment and maintenance of H3K27ac and 5hmC at enhancers and that this function can be independent of FOXA2. Whether the two epigenetic alterations established and maintained by HNF4A occur concomitantly or are unrelated remains to be determined.

REFERENCES

1. Sladek FM, Zhong WM, Lai E, Darnell JE. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev. 1990;4:2353–65.
2. Li J, Ning G, Duncan SA. Mammalian hepatocyte differentiation requires the transcription factor HNF-4alpha Genes Dev. 2000;14:464–474.
3. DeLaForest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA, et al. HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. Development. 2011;138:4143–53.
4. Bonzo JA, Ferry CH, Matsubara T, Kim J-H, Gonzalez FJ. Suppression of Hepatocyte Proliferation by Hepatocyte Nuclear Factor 4α in Adult Mice. J. Biol. Chem. 2012;287:7345–7356.
5. Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol. Cell. Biol. 2001;21:1393–403.
6. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, et al. Mutations in the hepatocyte nuclear factor-4α gene in maturity-onset diabetes of the young (MODY1). Nature. 1996;384:458–460.

7. Ning B-F, Ding J, Yin C, Zhong W, Wu K, Zeng X, et al. Hepatocyte Nuclear Factor 4 Suppresses the Development of Hepatocellular Carcinoma. Cancer Res. 2010;70:7640–7651.

8. Santangelo L, Marchetti A, Cicchini C, Conigliaro A, Conti B, Mancone C, et al. The stable repression of mesenchymal program is required for hepatocyte identity: A novel role for hepatocyte nuclear factor 4α. Hepatology. 2011;53:2063–2074.

9. Walesky C, Gunewardena S, Terwilliger EF, Edwards G, Borude P, Apte U. Hepatocyte-specific deletion of hepatocyte nuclear factor-4α in adult mice results in increased hepatocyte proliferation. Am. J. Physiol. Liver Physiol. 2013;304:G26–G37.

10. Watt A, Garrison WD, Duncan SA. HNF4: A central regulator of hepatocyte differentiation and function. Hepatology. 2003;37:1249–1253.

11. Alder O, Cullum R, Lee S, Kan AC, Wei W, Yi Y, et al. Hippo signaling influences HNF4A and FOXA2 enhancer switching during hepatocyte differentiation. Cell Rep. 2014;9:261–271.

12. Mueller B, Mieczkowski J, Kundu S, Wang P, Sadreyev R, Tolstorukov MY, et al. Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. Genes Dev. 2017;31:451–462.

13. Zhu H, Wang G, Qian J. Transcription factors as readers and effectors of DNA methylation. Nat. Rev. Genet. 2016;17:551–565.

14. Calo E, Wysocka J. Modification of enhancer chromatin: what, how and why? 2013;Mol Cell. 2013;49:825-837

15. Robertson AG, Bilenky M, Tam A, Zhao Y, Zeng T, Thiessen N, et al. Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. Genome Res. 2008;18:1906–1917.

16. Boyes J, Bird A. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. EMBO J. 1992;11:327–33.

17. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science (80-.). 2017;356:eaaj2239.

18. Schübeler D, Lorincz MC, Cimbora DM, Telling A, Feng YQ, Bouhassira EE, et al. Genomic targeting of methylated DNA: influence of methylation on transcription, replication, chromatin structure, and histone acetylation. Mol. Cell. Biol. 2000;20:9103–12.

19. Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, et al. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev. Cell. 2014;29:102–11.

20. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature. 2011;473:394–397.

21. Perera A, Eisen D, Wagner M, Laube SK, Künzel AF, Koch S, et al. TET3 Is Recruited by REST for Context-Specific Hydroxymethylation and Induction of Gene Expression. Cell Rep. 2015;11:283–294.

22. Choi SH, Gearhart MD, Cui Z, Bosnakovski D, Kim M, Schennum N, et al. DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. Nucleic Acids Res. 2016;44:5161–5173.

23. Fujiki K, Shinoda A, Kano F, Sato R, Shirahige K, Murata M. PPARγ-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. Nat. Commun. 2013;4:2262.

24. Yang YA, Zhao JC, Fong K-W, Kim J, Li S, Song C, et al. FOXA1 potentiates lineage-

specific enhancer activation through modulating TET1 expression and function. Nucleic Acids Res. 2016;44:8153–64.

25. Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, et al. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature. 2011;475:386–389.

26. Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature. 2011;475:390–393.

27. Nakamori D, Akamine H, Takayama K, Sakurai F, Mizuguchi H. Direct conversion of human fibroblasts into hepatocyte-like cells by ATF5, PROX1, FOXA2, FOXA3, and HNF4A transduction. Sci. Rep. 2017;7:16675.

28. Kreamer BL, Staecker JL, Sawada N, Sattler GL, Hsia MT, Pitot HC. Use of a low-speed, iso-density percoll centrifugation method to increase the viability of isolated rat hepatocyte preparations. In Vitro Cell. Dev. Biol. 1986;22:201–11.

29. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, et al. Methylome analysis using MeDIP-seq with low DNA concentrations. Nat. Protoc. 2012;7:617–636.

30. Wederell ED, Bilenky M, Cullum R, Thiessen N, Dagpinar M, Delaney A, et al. Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. Nucleic Acids Res. 2008;36:4549–64.

31. Tanimizu N, Saito H, Mostov K, Miyajima A. Long-term culture of hepatic progenitors derived from mouse Dlk+ hepatoblasts. J. Cell Sci. 2004;117:6425–6434.

32. Hong YH, Varanasi US, Yang W, Leff T. AMP-activated Protein Kinase Regulates HNF4α Transcriptional Activity by Inhibiting Dimer Formation and Decreasing Protein Stability. J. Biol. Chem. 2003;278:27495–27501.

33. Li J, Wu X, Zhou Y, Lee M, Guo L, Han W, et al. Decoding the dynamic DNA methylation and hydroxymethylation landscapes in endodermal lineage intermediates during

pancreatic differentiation of hESC. Nucleic Acids Res. 2018;46:2883–2900.

34. Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, et al. The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. Mol. Cell. 2016;62:79–91.

35. Younesy H, Nielsen CB, Lorincz MC, Jones SJM, Karimi MM, Möller T. ChAsE: chromatin analysis and exploration tool. Bioinformatics. 2016;32:3324–3326.

36. Guan W, Guyot R, Samarut J, Flamant F, Wong J, Gauthier KC. Methylcytosine dioxygenase TET3 interacts with thyroid hormone nuclear receptors and stabilizes their association to chromatin. Proc. Natl. Acad. Sci. U. S. A. 2017;114:8229–8234.

37. Perera A, Eisen D, Wagner M, Laube SK, Künzel AF, Koch S, et al. TET3 Is Recruited by REST for Context-Specific Hydroxymethylation and Induction of Gene Expression. Cell Rep. 2015;11:283–294.

38. Ancey P-B, Ecsedi S, Lambert M-P, Talukdar FR, Cros M-P, Glaise D, et al. TET-Catalyzed 5-Hydroxymethylation Precedes HNF4A Promoter Choice during Differentiation of Bipotent Liver Progenitors. Stem cell reports. 2017;9:264–278.

39. Rollini P, Fournier RE. The HNF-4/HNF-1alpha transactivation cascade regulates gene activity and chromatin structure of the human serine protease inhibitor gene cluster at 14q32.1. Proc. Natl. Acad. Sci. U. S. A. 1999;96:10308–13.

40. Holloway MG, Miles GD, Dombkowski AA, Waxman DJ. Liver-specific hepatocyte nuclear factor-4alpha deficiency: greater impact on gene expression in male than in female mouse liver. Mol. Endocrinol. 2008;22:1274–86.

41. Yoshida E, Aratani S, Itou H, Miyagishi M, Takiguchi M, Osumu T, et al. Functional Association between CBP and HNF4 inTrans-activation. Biochem. Biophys. Res. Commun. 1997;241:664–669.

42. Eeckhoute J, Formstecher P, Laine B. Hepatocyte nuclear factor 4alpha enhances the hepatocyte nuclear factor 1alpha-mediated activation of transcription. Nucleic Acids Res. 2004;32:2586–93.

43. Wang J-C, Stafford JM, Granner DK. SRC-1 and GRIP1 coactivate transcription with hepatocyte nuclear factor 4. J. Biol. Chem. 1998;273:30847–30850.

44. Zhang Q, Lei X, Lu H. Alterations of Epigenetic Signatures in Hepatocyte Nuclear Factor 4a Deficient Mouse Liver Determined by Improved ChIP-qPCR and (h)MeDIP-qPCR Assays. 2014;

45. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324:930–5.

46. Reizel Y, Sabag O, Skversky Y, Spiro A, Steinberg B, Bernstein D, et al. Postnatal DNA demethylation and its role in tissue maturation. Nat. Commun. 2018;9:2040.

47. Jin S-G, Zhang Z-M, Dunwell TL, Harter MR, Wu X, Johnson J, et al. Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. Cell Rep. 2016;14:493–505.

FIGURE LEGENDS

Fig.1 Active enhancer signatures are associated with HNF4A bound regions (HBRs). (A) Heatmap displaying the histone (H3K4me1, H3K4me3, and H3K27ac) and DNA modifications (5mC and 5hmC) patterns at HBRs in mouse hepatocytes. A 5kb window is shown with the summit of HNF4A binding in the centre of each panel. The coloured bar with blue, yellow and red colour represents low, moderate and high enrichment respectively (B) Aggregation plots of MeDIP (red) and hMeDIP (blue) signals at HBRs (black line). (C and D) UCSC genome browser view of *Cdc42bpb* and *Rfx4* loci showing enrichment of HNF4A (black), H3K27ac (orange), H3K4me1 (dark green), H3K4me3 (light green), 5mC (red), 5hmC (blue) and DNase-seq signals (pink) at HBR (highlighted). (E) Bar graph showing 5mC (Red bars) and 5hmC (blue) levels at HBR associated genes *Cdc42bpb, Ido2, Mgst3, Rfx4 and Tet3* (black line). *Dazl*, *Tbx15* and *IGd* genomic regions were used as 5mC positive, negative and non-CpG controls, respectively. (n=3). Values represent mean ±SEM.

Fig. 2 HNF4A bound regions show high H3K27ac and 5hmC in the absence of FOXA2. (A) Heatmap of K-means clustering (K=5) of TF (HNF4A and FOXA2) bound regions and H3K4me1, H3K4me3, H3K27ac, 5hmC and 5mC. Clusters Cl1 and Cl5 represent HNF4A only and FOXA2 only regions, respectively. Cl2/3/4 represents HNF4A/FOXA2 shared regions. A 5kb window is shown with transcription factor binding in the middle of each panel (B) Enrichment plots of H3K4me1 (C) H3K27ac (D) 5hmC (E) 5mC for Cl1, Cl2, Cl3, Cl4 and Cl5 represented by red, yellow, cyan, blue and black lines, respectively (F) Scatter dot plot representing RNA-seq Log2 RPKM values of all genes associated with clusters. A significant difference was tested using Kruskal-Wallis one-way ANOVA with Dunn's post-test correction; ****$p < 0.001$.

Fig. 3 HNF4A is required for maintenance of H3K27ac and 5hmC at enhancers. (A) Genomic map displaying ChIP (TF and histone modifications), MeDIP, hMeDIP and DNase-seq signals at HNF4A and FOXA2 shared sites at *Ido2* locus (highlighted region) (B) Example of HNF4A only bound region at *Acss3* locus and enrichment of histone and DNA modifications at HBR (highlighted region) (C) ChIP-qPCR displaying H3K27ac enrichment in WT and HNF4A conditional knockout (HNF4A $^{F/F}$; AlbERT2cre) livers at HNF4A/FOXA2 shared (*Apoa5, Nnmt, Agxt, Prlr and Rfx4*) and HNF4A only regions (*Hnmt, Ccnd1, Acss3 and G6pc*). *Gapdh* and *Sixl1* represent H3K27ac positive and negative control for WT liver samples (D) Bar plot showing the 5hmC level at HNF4A/FOXA2 shared (*Apoa5, Nnmt, Agxt, Prlr and Rfx4*) and HNF4A only genes (*Hnmt, Ccnd1, Acss3 and G6pc*)

Fig. 4 Gain of HNF4A binding during hepatoblast to hepatocyte differentiation accompanies histone and DNA modifications. (A) Genome maps for HNF4A target genes *Apoa5* and *Nnmt*, with normalised ChIP-seq signals for H3K4me1 (dark green), H3K27ac (orange) and HNF4A (black) from hepatoblast and hepatocyte showing a gain of HNF4A and increased level of H3K27ac at hepatocyte-specific HBR (highlighted). DNase-seq data from the liver is shown in pink. (B) Heatmap showing HNF4A sites gained during hepatoblast to hepatocyte differentiation are accompanied by H3K27ac. Colour bar below heat map represents the enrichment level. (C) 5mC level at HNF4A gained loci in hepatoblasts and hepatocytes. (D) 5hmC level at HNF4A gained sites in hepatoblasts and hepatocytes. (n=3). Values represent mean ±SEM **P*<0.05, ***P*<0.01 by *t*-test. (E) RNA-seq with log2RPKM values of HNF4A target genes (*Apoa5, Nnmt, Agxt, Prlr, Ido2* and *Rfx4*) in hepatoblasts and hepatocytes. (F) Box plot showing RNA-seq values (Log2RPKM) in hepatoblast and hepatocyte for genes that gained HNF4A binding during differentiation.

Fig. 5 HNF4A interacts with TET3 *in vivo* and *in vitro.* (A) RNA-Seq RPKM values of *Tet1*, *Tet2* and*Tet3* in hepatocytes. (B) Western blot showing expression of TET3 in hepatoblast and

hepatocytes with β-Actin control. (C) Immunoprecipitation and western analysis of HepG2 lysate which was subjected to immunoprecipitation using anti-HNF4A and negative control (IgG) antibody followed by western for HNF4A and TET3. (D) Hepatocytes isolated from 3 different female livers were subjected to immunoprecipitation using anti-HNF4A and negative control (IgG) antibody followed by western using HNF4A and TET3 antibody. (E) Immunoprecipitation and western analysis of hepatoblast lysates subjected to IP using anti-HNF4A and IgG antibody. (F) Cartoon describing the proximity ligation assay (PLA) technique (top). PLA on adult (middle) and embryonic liver (bottom) sections using an antibody against HNF4A and TET3 confirm endogenous interaction shown as yellow dots inside DAPI stained nuclei (Blue).

Fig. 6 *Tet3* transcription is directly regulated by HNF4A. (A) Genomic maps showing HNF4A, H3K4me1 and H3K27ac enrichment at *Tet3* locus in hepatoblast (highlighted) (B) *Hnf4a* and *Tet3* mRNA and protein level in siControl and *siHnf4a* knockdown in HPPL cells (96 hrs). (C) Luciferase assays using control (PGL3-EIBP-Empty and PDGT-Empty), HNF4A overexpression (PDGT-HNF4A) and the mouse *Tet3* enhancer (PGL3EIBP-mTet3E) constructs. Constructs were overexpressed in HEK293T cells in combinations shown below the graph (+ transfected and – not transfected) (D) Luciferase assays in HepG2 cells transfected with control (PGL3-EIBP-empty) or *Tet3* enhancer luciferase constructs (PGL3EIBP-mTET3E). (E) Schematic representation of the *Tet3* enhancer region (mTet3E) cloned for luciferase assays. Sequence alignments show two HNF4A motifs (HNFA-1 and HNF4A-2) within the *Tet3* enhancer. The table represents the constructs generated by deleting HNF4A motifs. (F) Luciferase assays in HepG2 cells transfected with mTET3E or mTET3E carrying one deleted HNF4A motif (mTET3EΔ1 or mTET3EΔ2) or mTET3E with both HNF4A motifs deleted (mTET3Δ1Δ2). Values represent the mean ± SEM. p*<0.05, **p<0.01, ****p<0.0001 from *t-test*.

Fig. 7 Role of HNF4A in establishing and maintaining epigenetic modifications at enhancers. HNF4A is required for the maintenance of H3K27ac and 5hmC in mouse (top) and human (bottom). Coloured bars represent the TF (HNF4A- black) occupancy and level of H3K27ac (orange), 5mC (red) and 5hmC (blue) in WT and HNF4A KO liver. Loss of H3K27ac and 5hmC at gene regulatory regions result in decreased expression (sky-blue) from HNF4A target genes. Cluster 1 (left panel) represents HNF4A only regions where HNF4A interacts with CBP/p300 and TET3 to maintain H3K27ac and 5hmC levels at HBRs. Loss of HNF4A in mouse liver (HNF4A KO liver) shows reduction in H3K27ac and 5hmC and gain of 5mC at cluster 1 regions, represented by coloured bars. Cluster 2, 3 and 4 (middle panel) represents HNF4A and FOXA2 shared regions where 5hmC and H3K27ac level was maintained by HNF4A's interaction with epigenetic modifiers such as TET3 and CBP/p300. Similar to the cluster 1, the loss of HNF4A results in reduction of H3K27ac, 5hmC and gain of 5mC which directly affects transcription from associated genes. Cluster 5 (right panel) represents regions only occupied by FOXA2 in hepatocytes, with very low H3K27ac and 5hmC but having higher 5mC (right panel). The genes associated with these regions show lower expression (sky blue bars) as compared to the other clusters. In human liver carcinoma cells (HepG2 cells), HNF4A KD also produced similar changes in the level of 5hmC and 5mC (shown below the three mouse panels).

SUPPLEMENTARY FIGURE LEGENDS

Fig. S1 Epigenetic signatures associated with HNF4A binding in hepatocytes. (A) Genomic distribution of HNF4A bound regions in hepatocytes (B) DNase-seq signals at all HBRs in mouse liver (C and D) UCSC genome browser view showing enrichment of HNF4A, H3K27ac H3K4me1, H3K4me3, 5mC, 5hmC and DNase-seq signals at HBR (highlighted). (E and F) UCSC maps of *Dazl* and *Tbx15* displaying the level of 5mC at respective CpG Islands.

Fig. S2 HNF4A does not require a FOXA family member for its influence on epigenetic modifications at HBRs. (A) DNase-seq signals for each cluster in hepatocytes (B) Motif analysis of Cluster 1 regions showing enrichment for potential transcription factor binding sites. (C) Motif analysis of Cluster 5 regions showing enrichment for potential transcription factor binding sites. (D) Pathway analysis of cluster 1 associated genes in liver. (E) Pathway analysis of cluster 5 genes in liver.

Fig. S3 HNF4A maintains a high level of 5hmC at its binding sites *in vitro* (A) The 5mC level at shared and HNF4A only gene loci in WT and HNF4A KO livers. n=3.Values represent mean ±SEM. (B) Left: qRT-PCR analysis of HNF4A in HepG2 cells 72 hours following HNF4A knockdown (siHNF4a) to control (siCntrl) data is normalised to GAPDH. (n=3). Values represent the mean ± SEM. **p<0.01 from *t-test*. Right: WB analysis of HNF4A in HepG2 cells 72 hours following HNF4A knockdown (siHNF4a). (C) ChIP-qPCR analysis displaying enrichment of HNF4A at *RBKS*, *MGST2*, *SLCO4C1*, APOA4, *PRLR* and *AGXT* loci. (n=3) Values represent the mean ± SEM. *DAZL* promoter and region on *Chr17* serve as negative control. (D) hMeDIP-qPCR of HNF4A occupied regions at *RBKS*, *MGST2*, *SLCO4C1*, *PRLR* and *AGXT* in siCntrl and siHNF4A HepG2 samples. (E) MeDIP qPCR of HNF4A target genes in HepG2, *DAZL* serves as a 5mC positive control.

Fig. S4 Hepatoblast-enriched HNF4A binding shows decreased H3K27ac and reduced transcription from associated genes in hepatocytes (A) heat map displaying HNF4A and H3K27ac ChIP signals in hepatoblasts and hepatocytes. Colour bar below indicates the strength of signals. (B) Box plot showing RNA-seq values (Log2RPKM) in hepatoblasts and hepatocytes for genes that lost HNF4A binding during differentiation. Significance was tested using a *t*-test, ****p<0.0001.

Fig. S5 HNF4A and TET3 interact to regulate HNF4A target genes in hepatic cells. (A) Western showing protein expression of TET2 in hepatoblasts and hepatocytes (top) and β-Actin as loading control (bottom). (B) Co-immunoprecipitation in which Myc-HNF4A and flag-TET3 were ectopically expressed in 293T cells. IP was performed using anti-Myc and a negative control antibody (IgG), followed by western analysis with Flag (top) and Myc (bottom) antibody. (C) mRNA levels of *TETs* in HepG2 cells. (D) Proximity Ligation Assay for ectopically expressed Myc-HNF4A and Flag-TET3 confirming the interaction in HEK293T cells (shown by Red dots). (E) Secondary only probes for PLA analysis for embryonic (left) and adult liver (right) sections. Nuclei are stained with DAPI (blue).

Fig. S6 HNF4A regulates TET3 expression (A) Luciferase assays were performed in HEK293T cells. Cells were transfected with mouse *Tet3* enhancer construct (PGL3EIBP-mTET3E) or enhancer carrying one deleted HNF4A motif (mTET3EΔ1 or mTET3EΔ2) and both deleted HNF4A motifs (mTET3Δ1Δ2) with overexpression of HNF4A (PDGT-HNF4A). (B) UCSC maps using ENCODE ChIP-seq data displaying HNF4A, H3K4me1 and H3K27ac enrichment at *TET3* locus in HepG2 cells. The highlighted region was cloned in luciferase vector. (C) Luciferase assay using control (PGL3-EIBPEmpty and PDGT-empty), HNF4A overexpression (PDGT-HNF4A) and cloned human *TET3* (PGL3EIBP-hTET3) enhancer constructs. Constructs were overexpressed in HEK293T cells in combinations shown below the graph (+ transfected and – not transfected). (D) Luciferase assays in HepG2 cells transfected with control (PGL3-EIBP-empty) or human *TET3* (PGL3EIBP-hTET3E) enhancer luciferase constructs.

Fig. 1

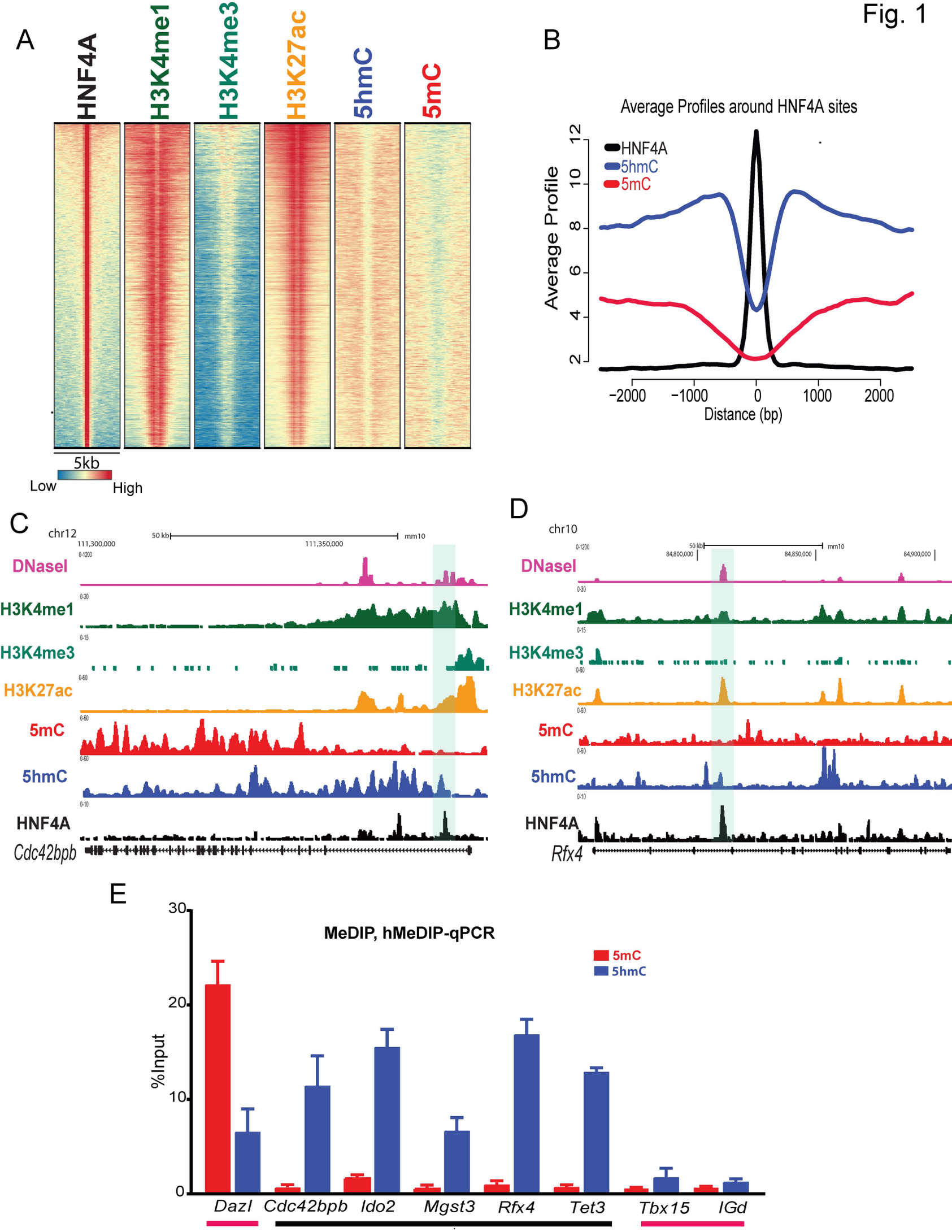
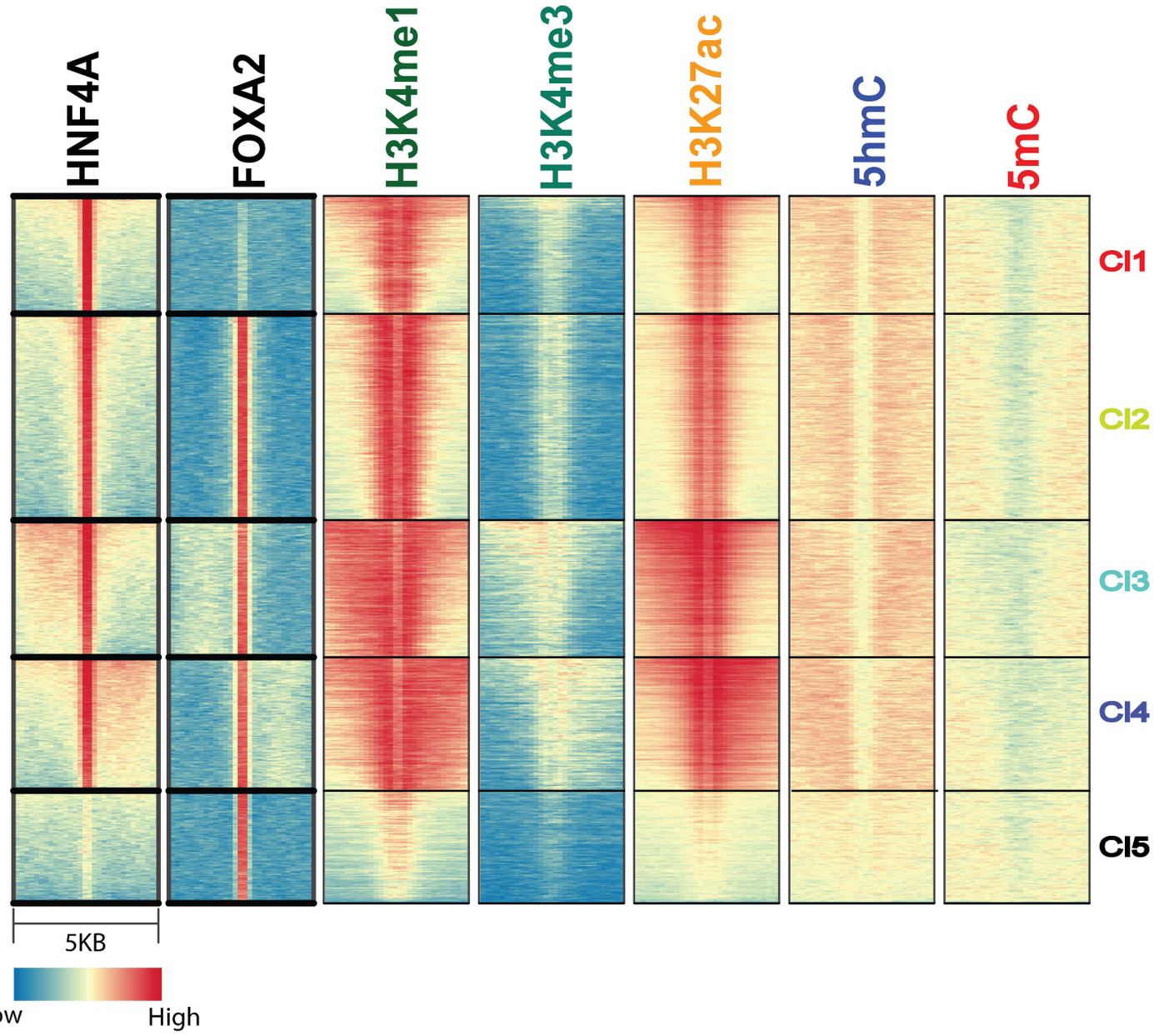
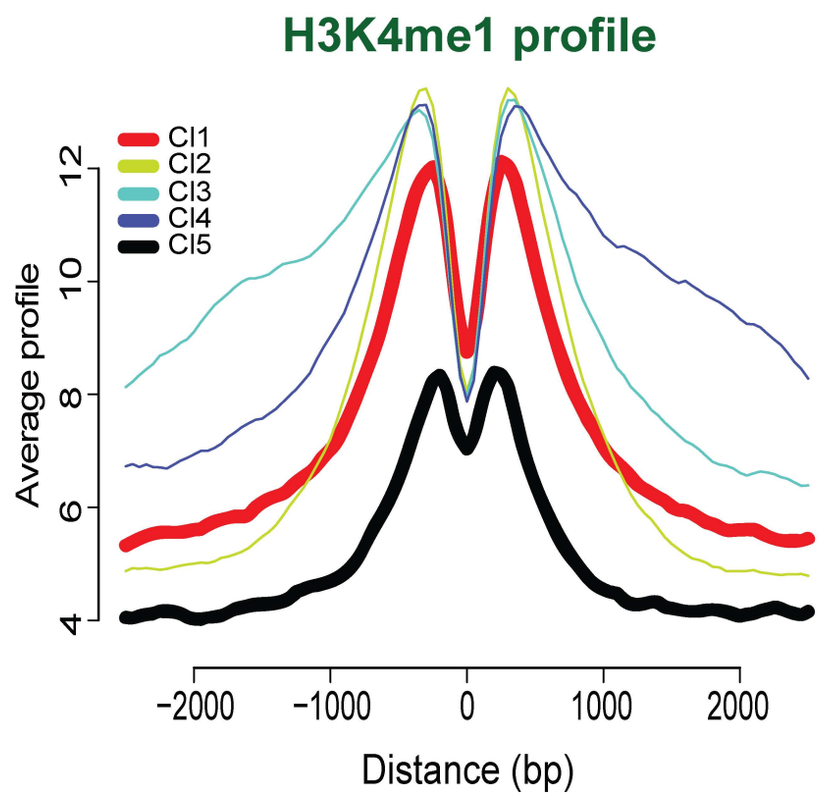

A
HNF4A
H3K4me1
H3K4me3
H3K27ac
5hmC
5mC
5kb
Low
High
B
Average Profiles around HNF4A sites
HNF4A
5hmC
5mC
Average Profile
Distance (bp)
C
chr12
DNaseI
H3K4me1
H3K4me3
H3K27ac
5mC
5hmC
HNF4A
Cdc42bpb
D
chr10
DNaseI
H3K4me1
H3K4me3
H3K27ac
5mC
5hmC
HNF4A
Rfx4
E
MeDIP, hMeDIP-qPCR
5mC
5hmC
%Input
Dazl
Cdc42bpb
Ido2
Mgst3
Rfx4
Tet3
Tbx15
IGd

Fig. 2

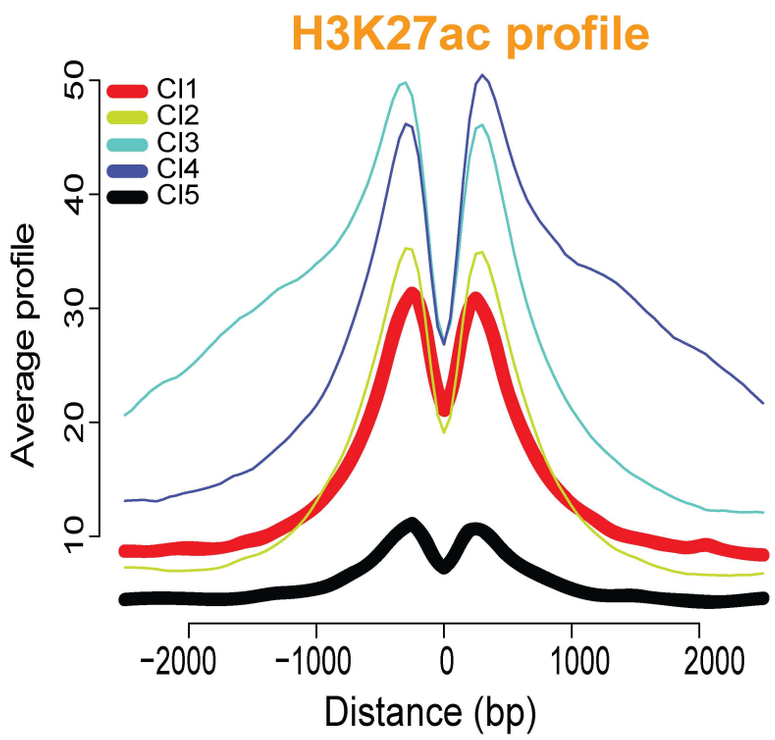
A


HNF4A
FOXA2
H3K4me1
H3K4me3
H3K27ac
5hmC
5mC
CI1
CI2
CI3
CI4
CI5
5KB
Low
High

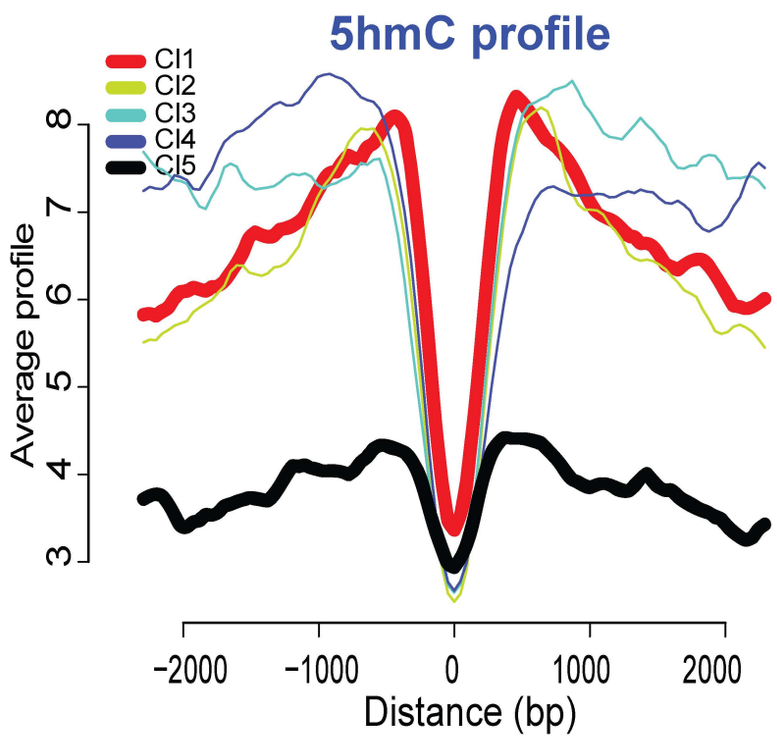
B


H3K4me1 profile
CI1
CI2
CI3
CI4
CI5
Average profile
4
6
8
10
12
−2000
−1000
0
1000
2000
Distance (bp)

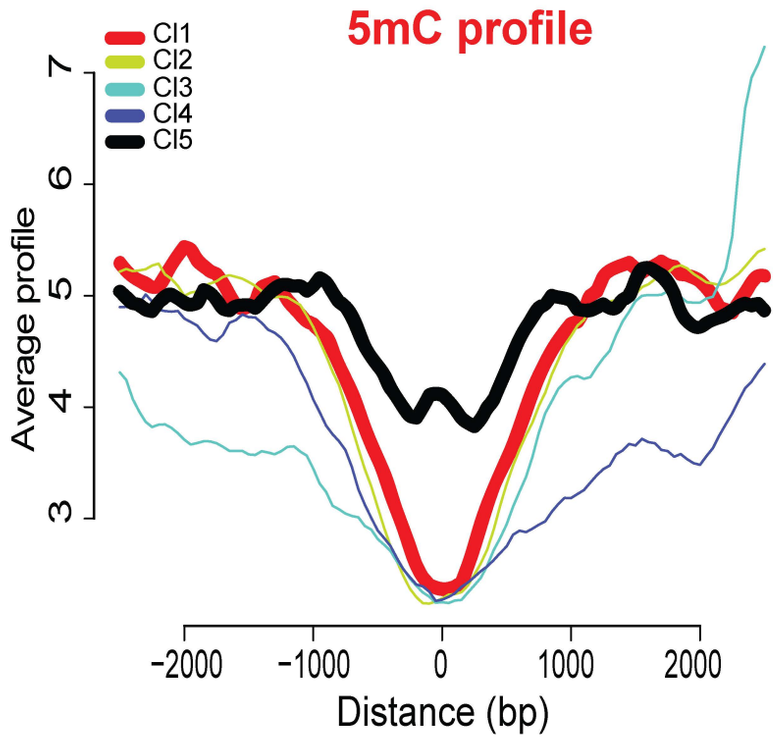
C


H3K27ac profile
CI1
CI2
CI3
CI4
CI5
Average profile
10
20
30
40
50
−2000
−1000
0
1000
2000
Distance (bp)

D


5hmC profile
CI1
CI2
CI3
CI4
CI5
Average profile
3
4
5
6
7
8
−2000
−1000
0
1000
2000
Distance (bp)

E


5mC profile
CI1
CI2
CI3
CI4
CI5
Average profile
3
4
5
6
7
−2000
−1000
0
1000
2000
Distance (bp)

F

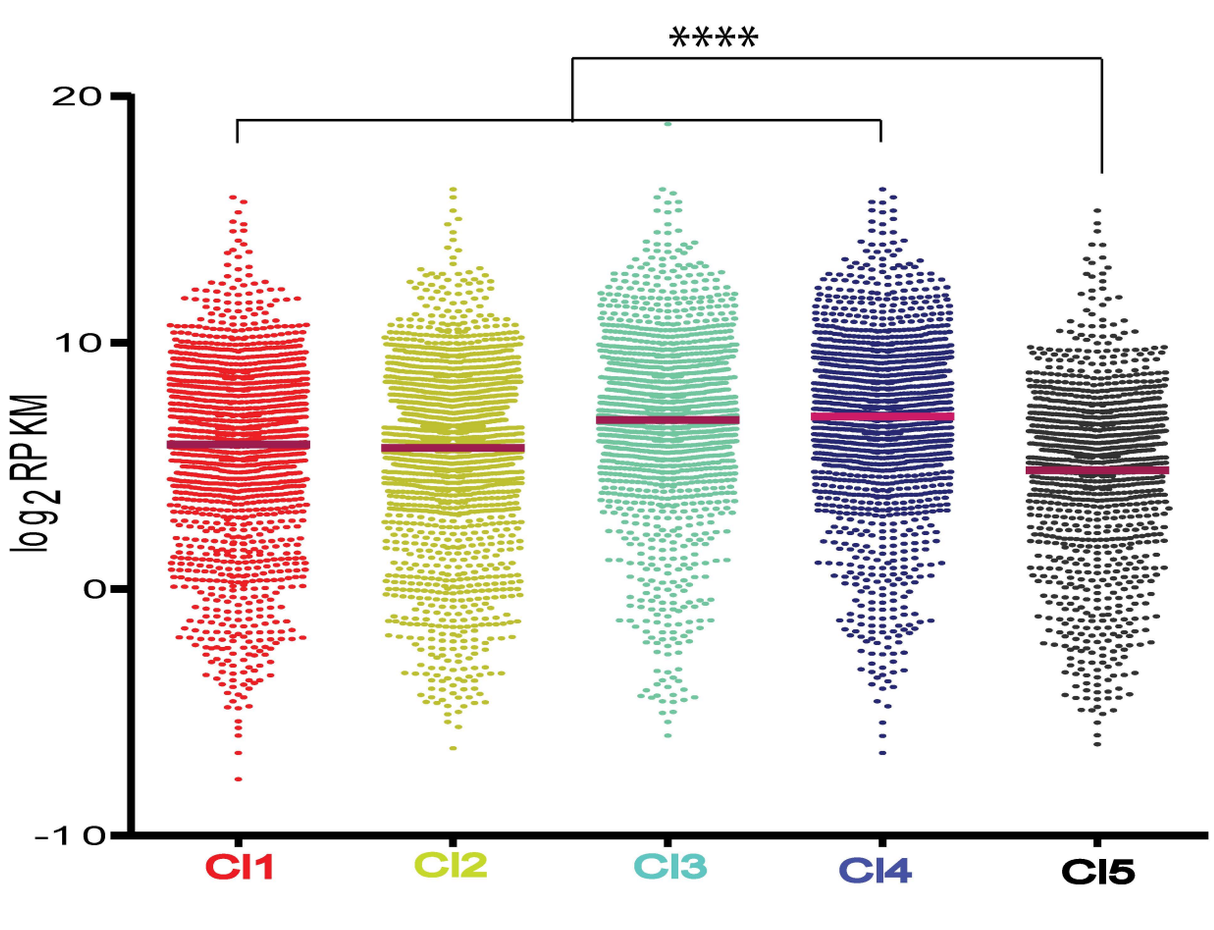
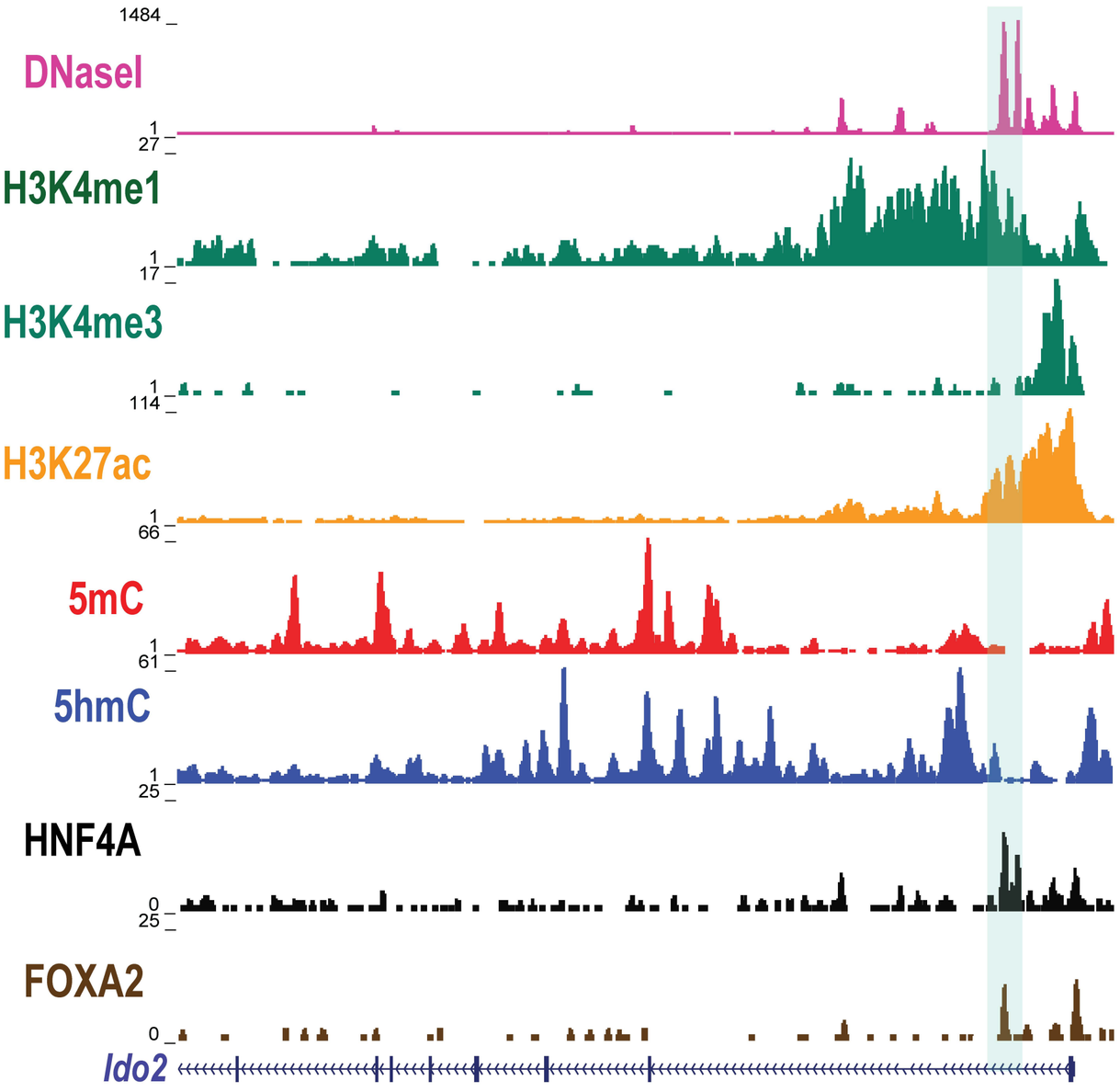
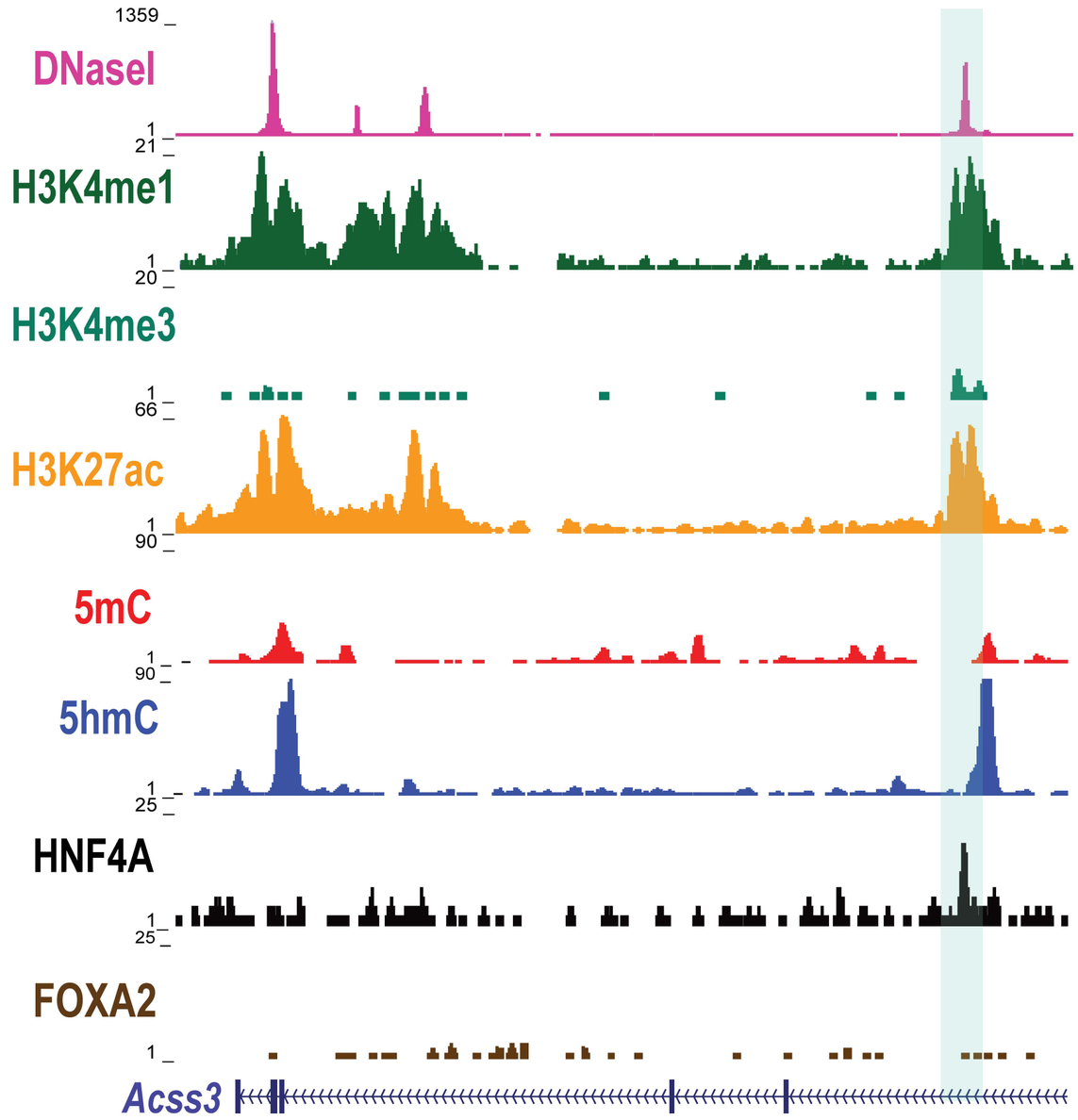

****
20
10
0
−10
log2 RPKM
CI1
CI2
CI3
CI4
CI5

Fig. 3

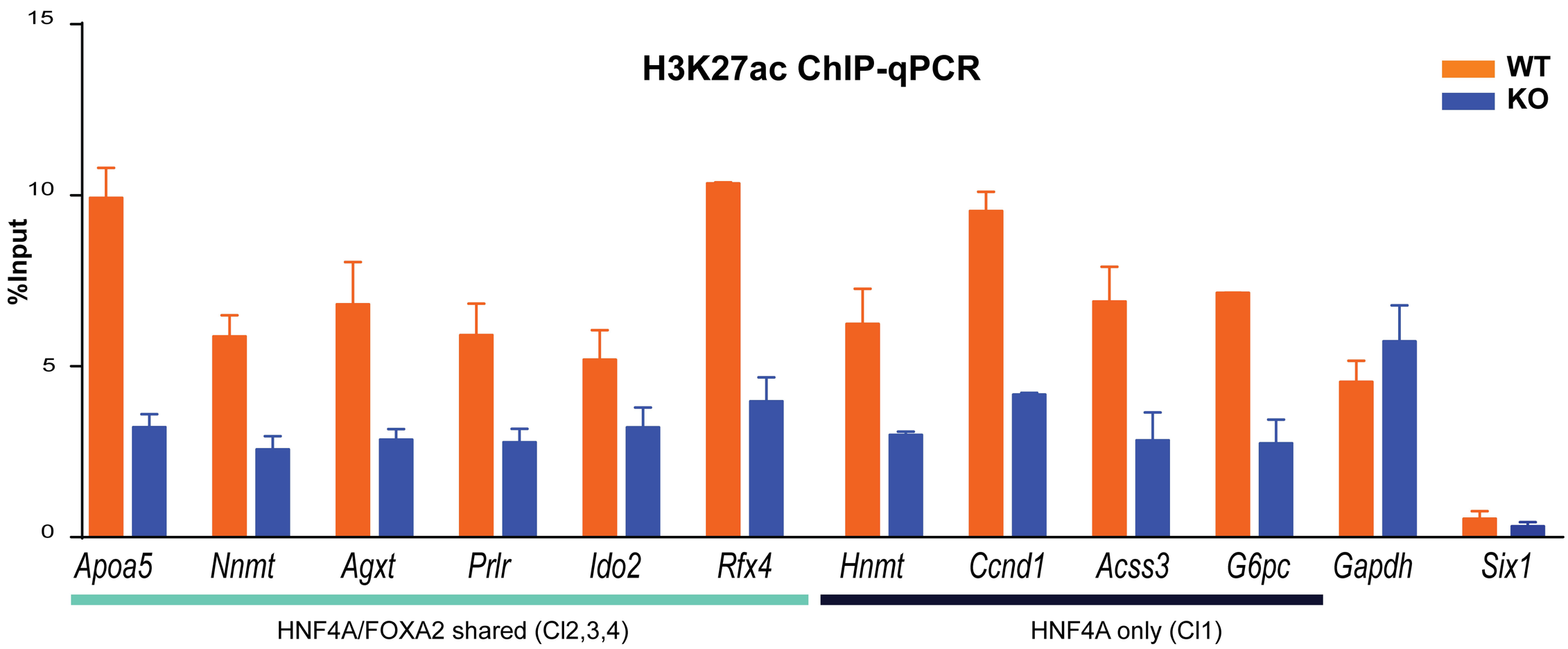
A


1484
DNaseI
1
27
H3K4me1
1
17
H3K4me3
1
114
H3K27ac
1
66
5mC
1
61
5hmC
1
25
HNF4A
0
25
FOXA2
0
Ido2

B


1359
DNaseI
1
21
H3K4me1
1
20
H3K4me3
1
66
H3K27ac
1
90
5mC
1
90
5hmC
1
25
HNF4A
1
25
FOXA2
1
Acss3

C


H3K27ac ChIP-qPCR
WT
KO
%Input
15
10
5
0
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4
Hnmt
Ccnd1
Acss3
G6pc
Gapdh
Six1
HNF4A/FOXA2 shared (Cl2,3,4)
HNF4A only (Cl1)

D

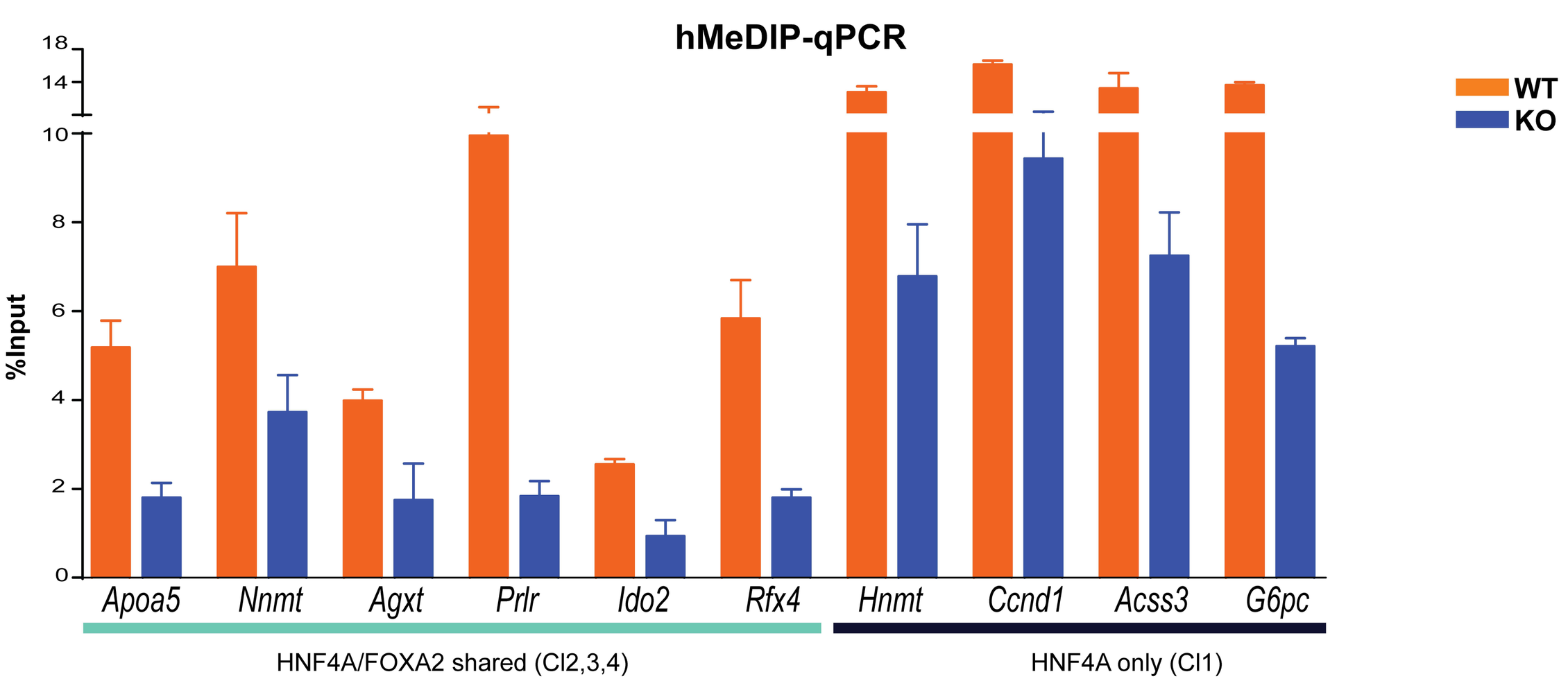
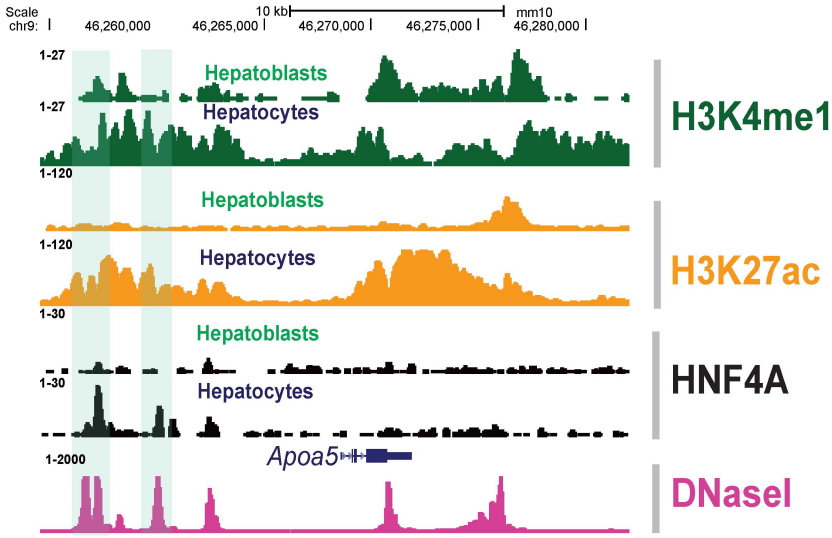
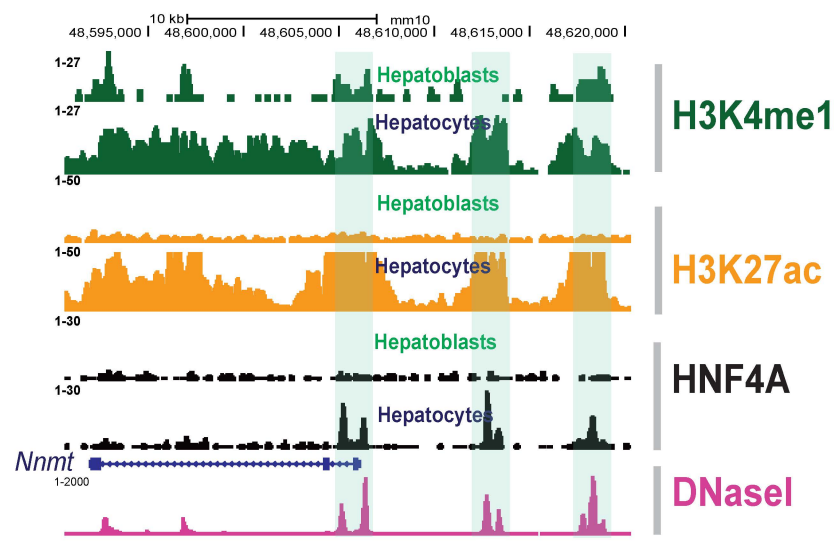

hMeDIP-qPCR
WT
KO
%Input
18
14
10
8
6
4
2
0
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4
Hnmt
Ccnd1
Acss3
G6pc
HNF4A/FOXA2 shared (Cl2,3,4)
HNF4A only (Cl1)

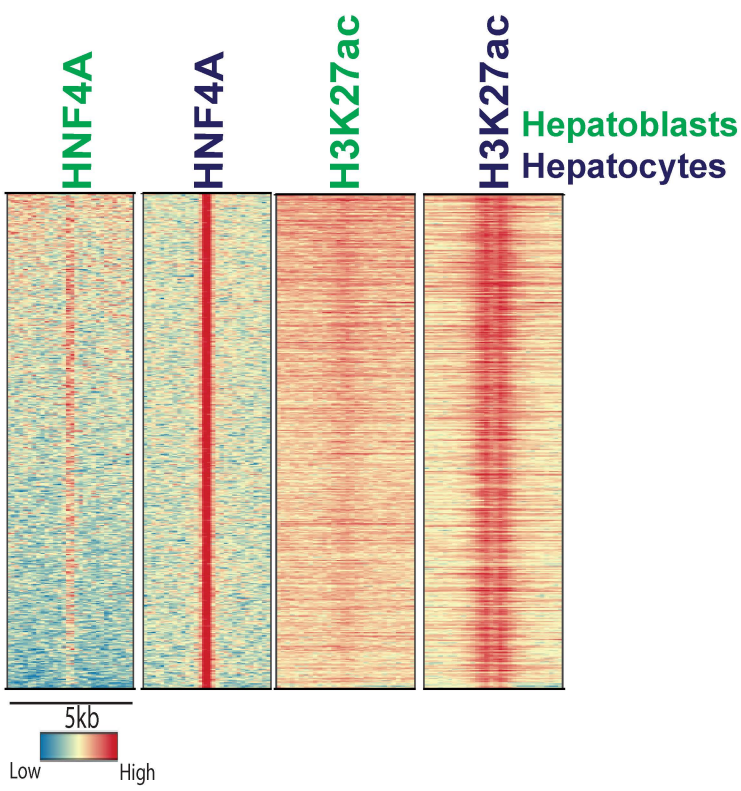
Fig. 4

A

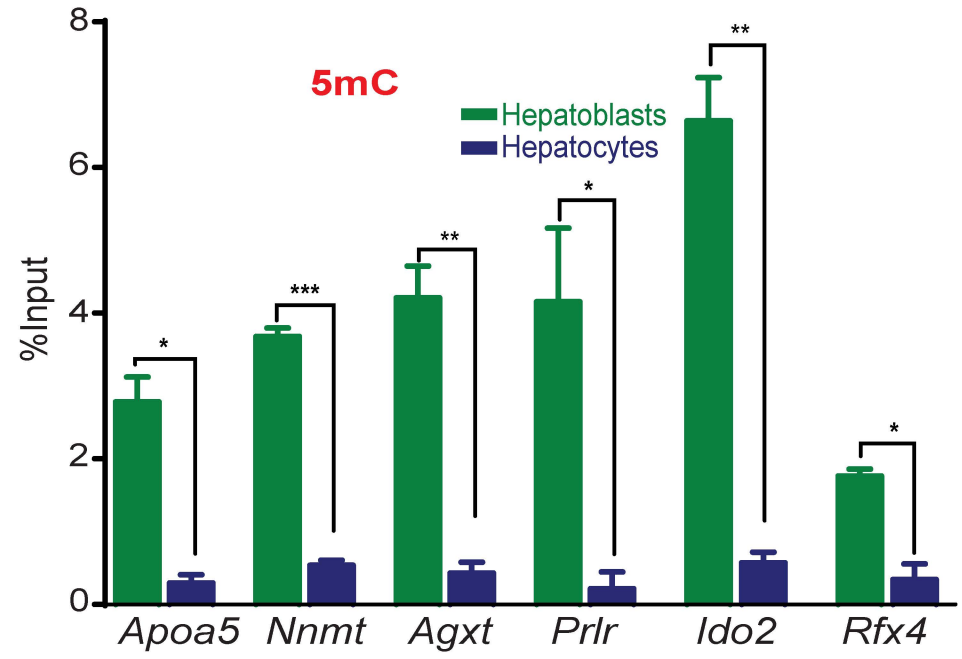

Scale
chr9:
46,260,000
10 kb
46,265,000
46,270,000
46,275,000
mm10
46,280,000
Hepatoblasts
Hepatocytes
H3K4me1
H3K27ac
HNF4A
Apoa5
DNaseI


10 kb
48,595,000
48,600,000
48,605,000
48,610,000
mm10
48,615,000
48,620,000
Hepatoblasts
Hepatocytes
H3K4me1
H3K27ac
HNF4A
Nnmt
DNaseI

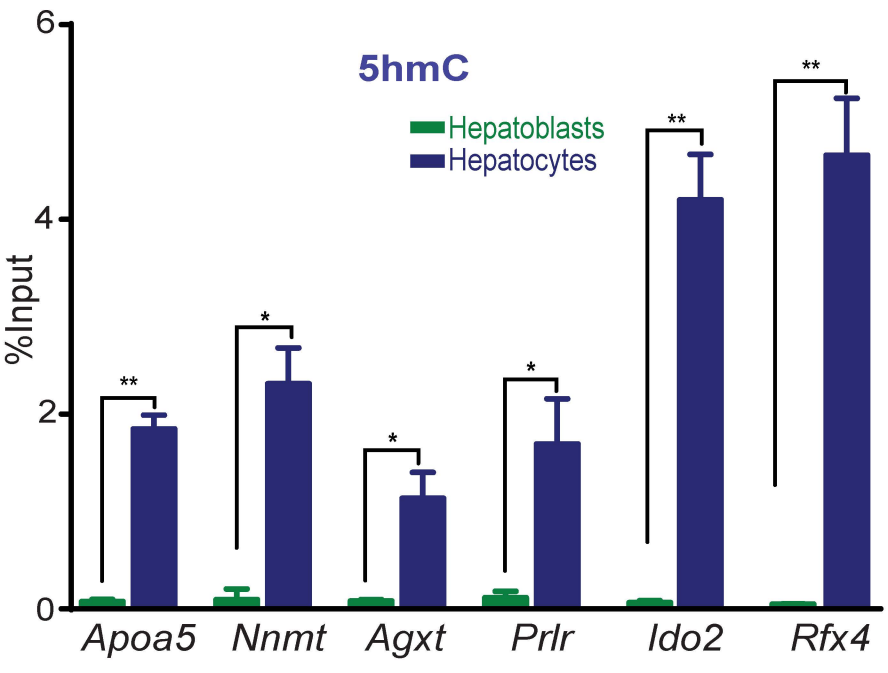
B


HNF4A
HNF4A
H3K27ac
H3K27ac
Hepatoblasts
Hepatocytes
5kb
Low
High

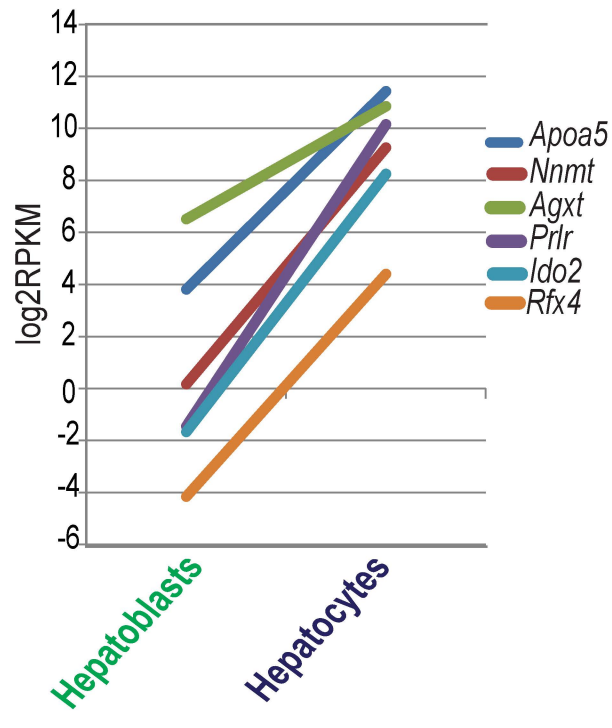
C


5mC
Hepatoblasts
Hepatocytes
%Input
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4

D


5hmC
Hepatoblasts
Hepatocytes
%Input
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4

E


log2RPKM
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4
Hepatoblasts
Hepatocytes

F

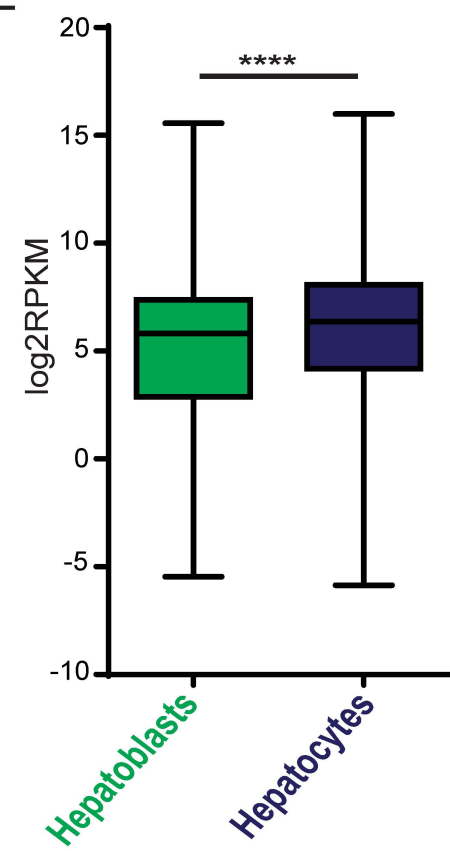
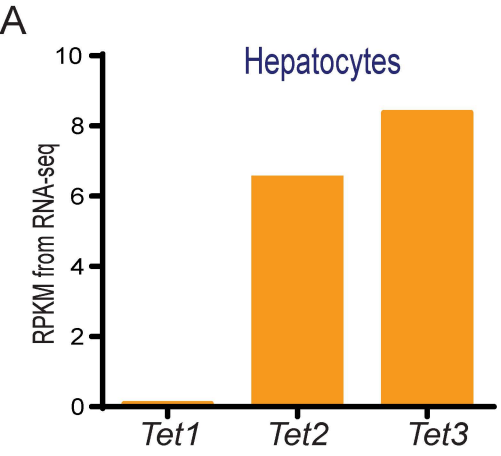
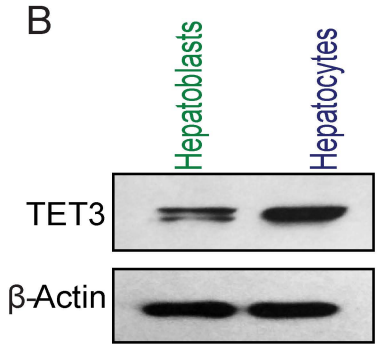
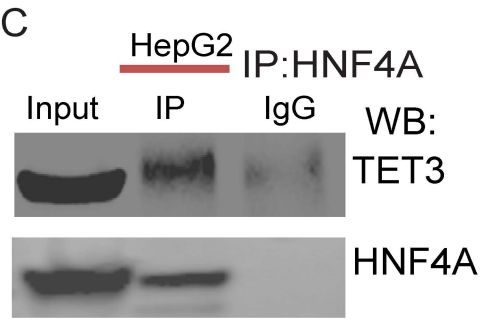
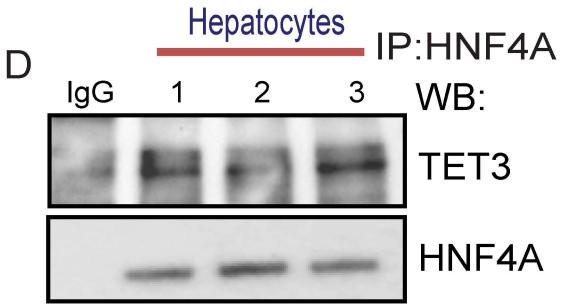

****
log2RPKM
Hepatoblasts
Hepatocytes

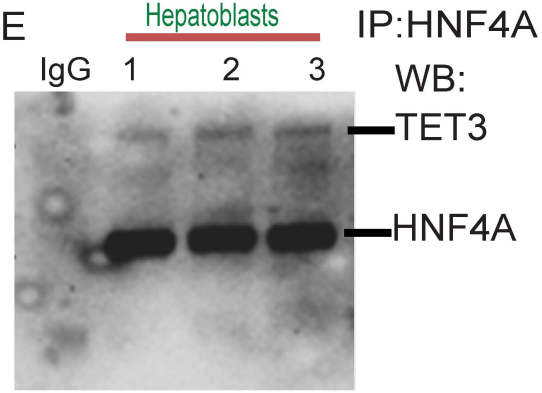
Fig. 5

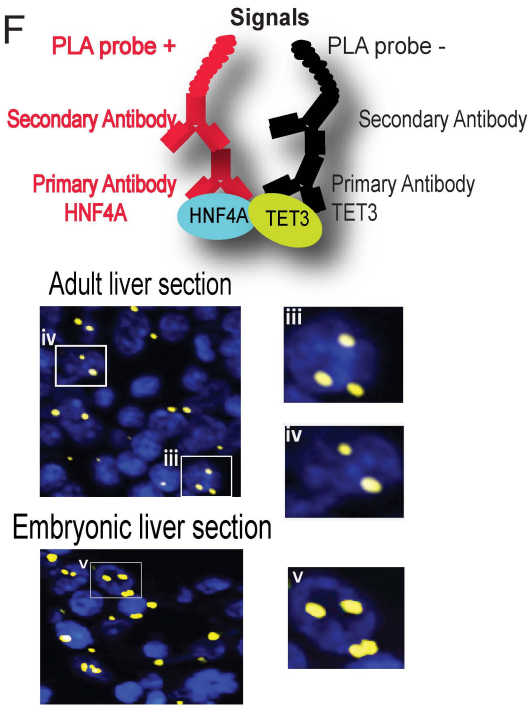

A
Hepatocytes
RPKM from RNA-seq
10
8
6
4
2
0
Tet1
Tet2
Tet3


B
Hepatoblasts
Hepatocytes
TET3
β-Actin

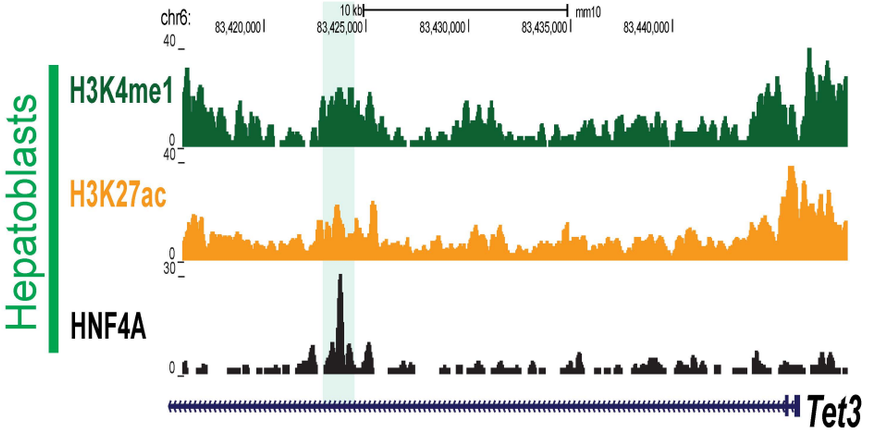

C
HepG2
IP:HNF4A
Input
IP
IgG
WB:
TET3
HNF4A


D
Hepatocytes
IP:HNF4A
IgG
1
2
3
WB:
TET3
HNF4A

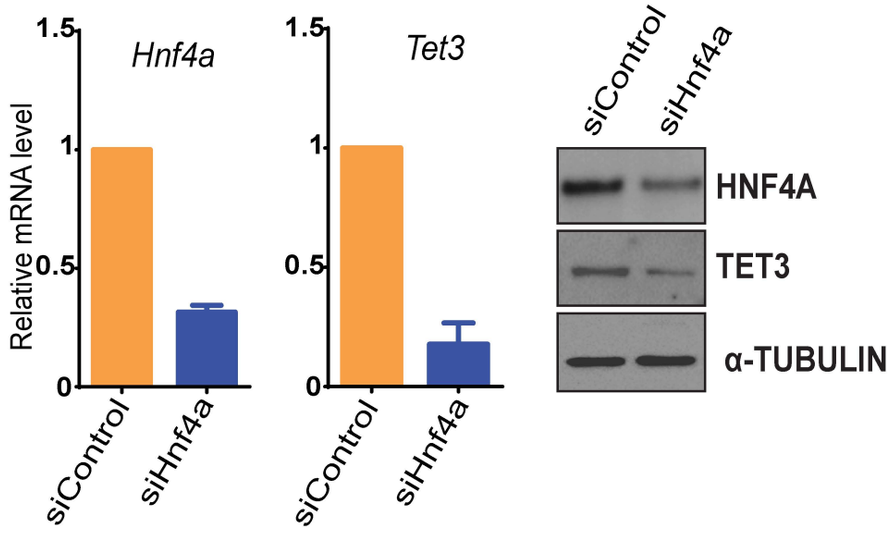

E
Hepatoblasts
IP:HNF4A
IgG
1
2
3
WB:
TET3
HNF4A


F
Signals
PLA probe +
PLA probe -
Secondary Antibody
Secondary Antibody
Primary Antibody
HNF4A
Primary Antibody
TET3
HNF4A
TET3
Adult liver section
iv
iii
iii
iv
Embryonic liver section
v
v

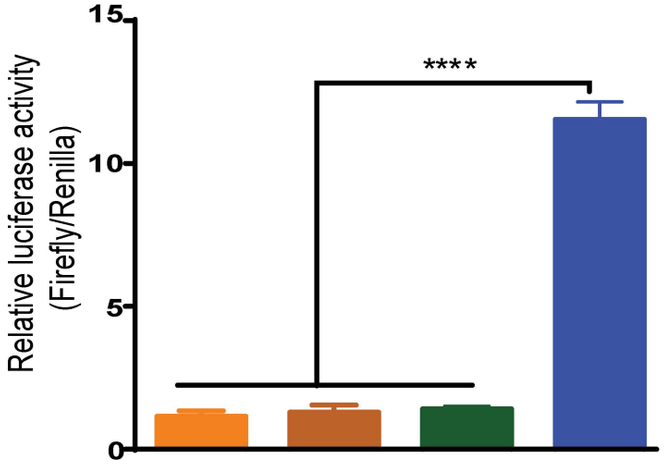
A



B

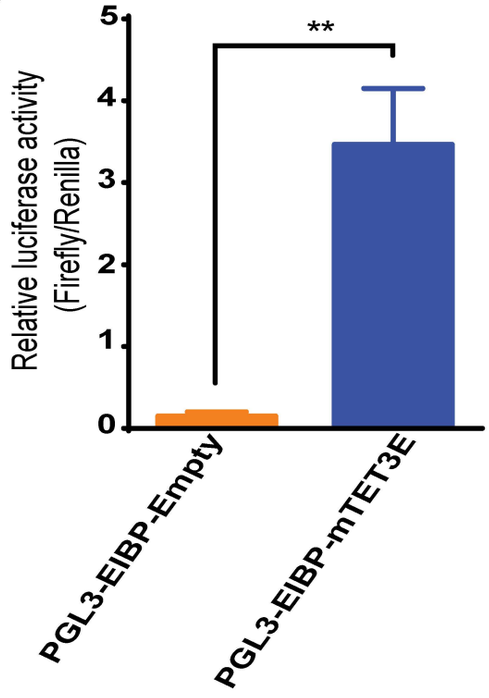


C

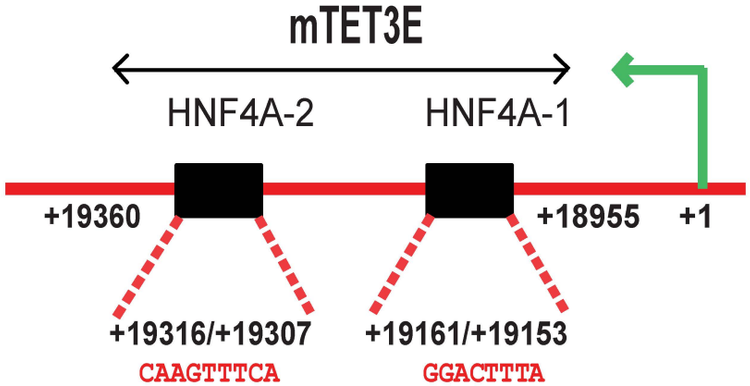


PGL3-EIBP-Empty	+	-	+	-
PGL3-EIBP-mTET3E	-	+	-	+
PDGT-Empty	+	+	-	-
PDGT-HNF4A	-	-	+	+

D



E



mTET3E Δ1	Present	Deleted
mTET3E Δ2	Deleted	Present
mTET3E Δ1Δ2	Deleted	Deleted

F

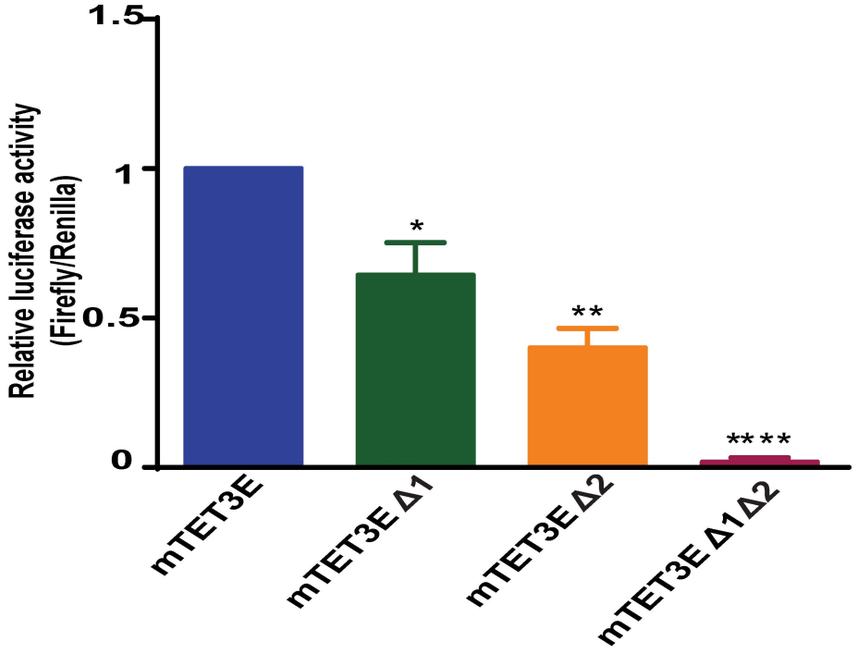


Fig. 7

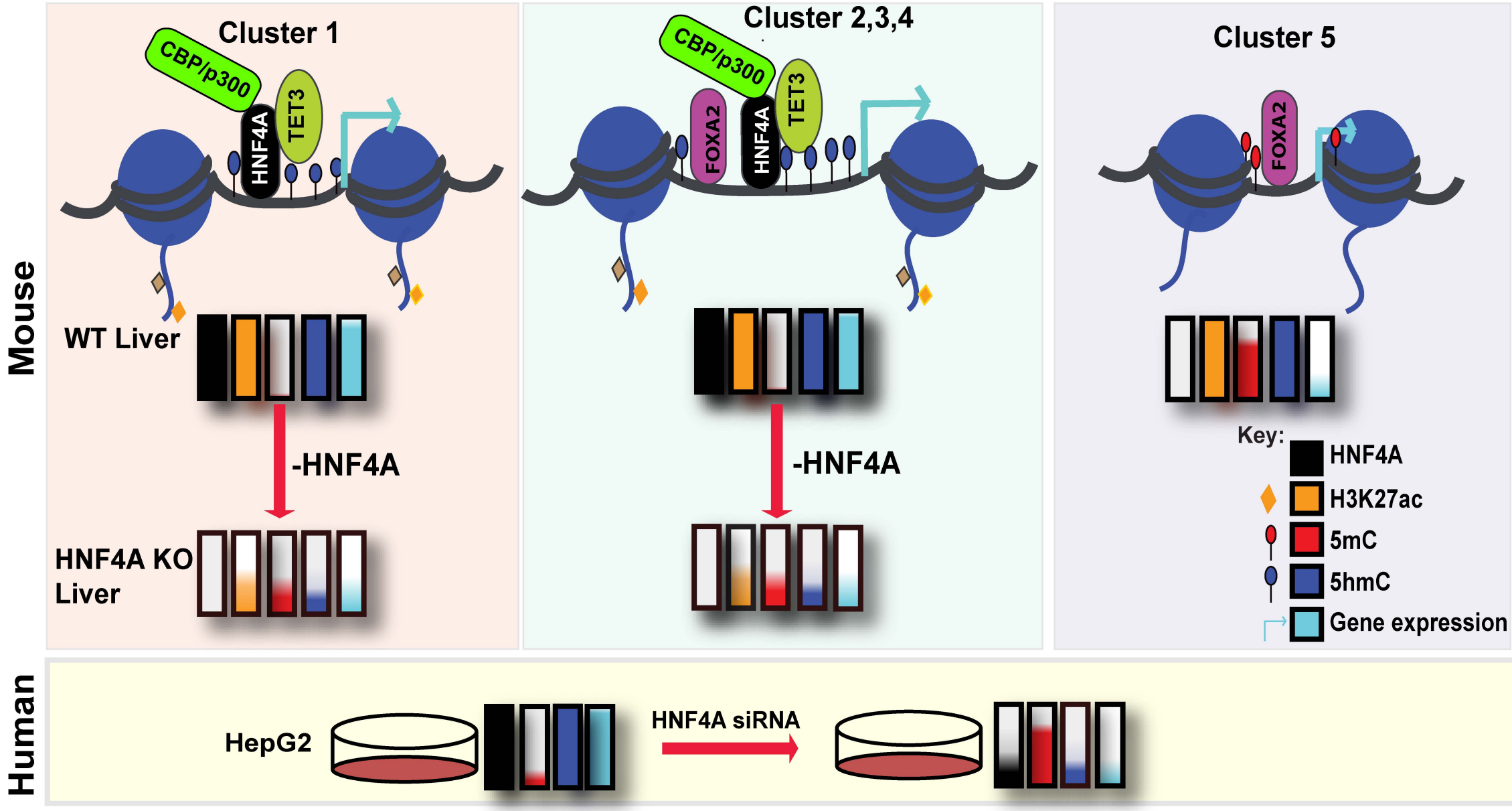
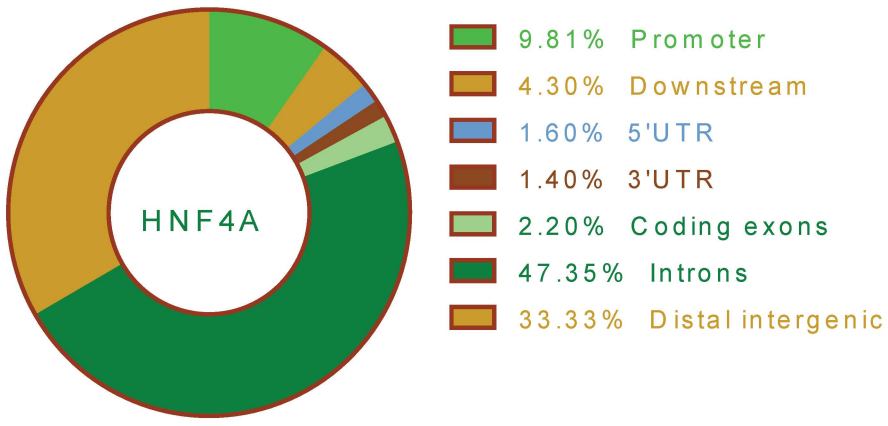
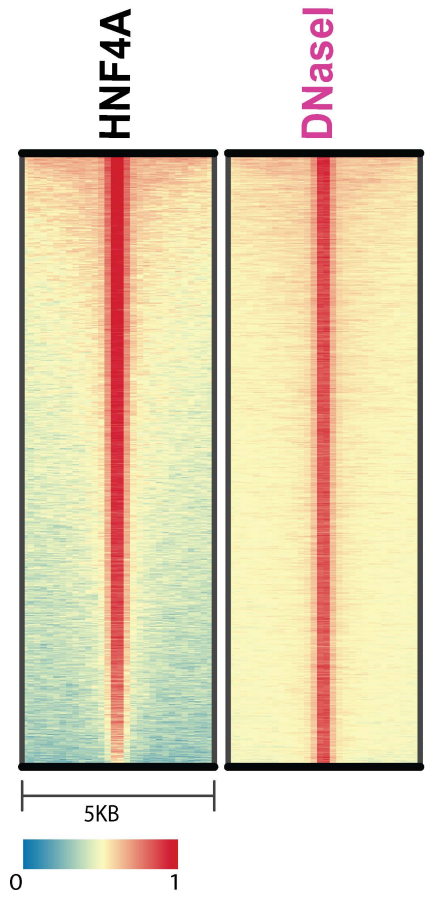

Mouse
Cluster 1
CBP/p300
HNF4A
TET3
WT Liver
-HNF4A
HNF4A KO Liver
Cluster 2,3,4
CBP/p300
FOXA2
HNF4A
TET3
-HNF4A
Cluster 5
FOXA2
Key:
HNF4A
H3K27ac
5mC
5hmC
Gene expression
Human
HepG2
HNF4A siRNA

Fig. S1

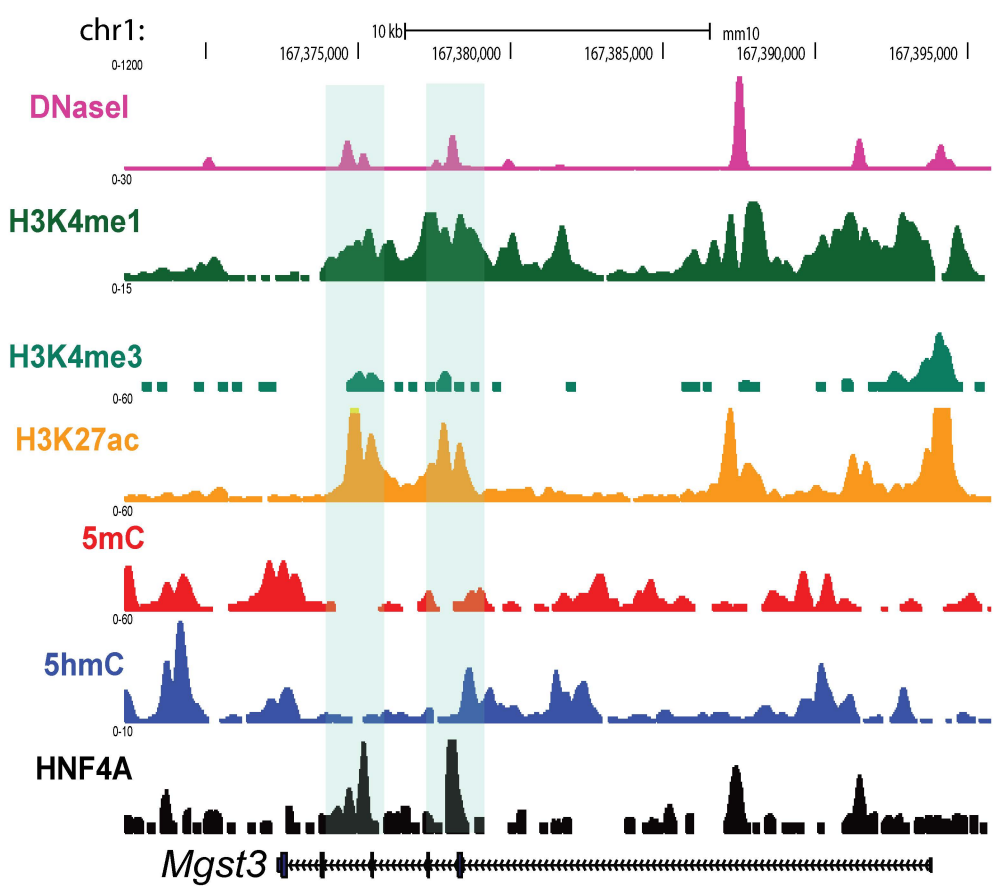
A



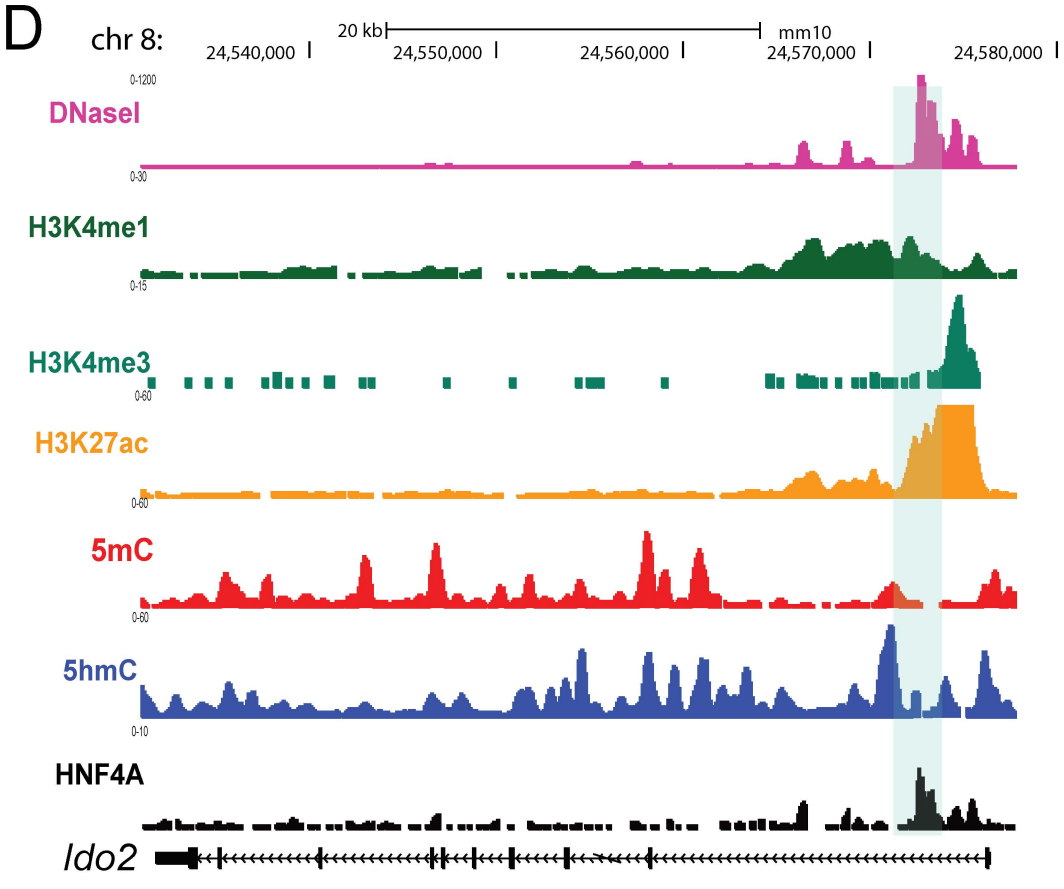
B



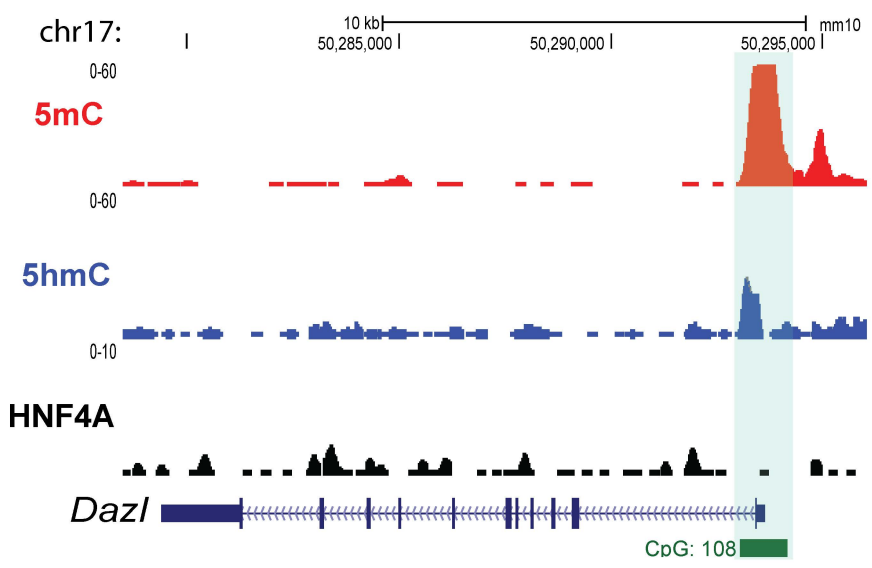
C



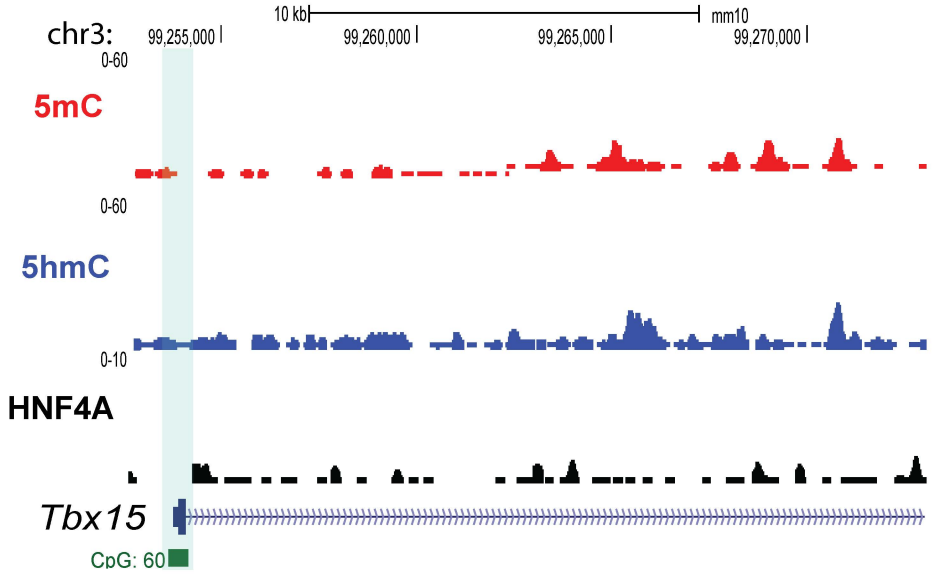
D



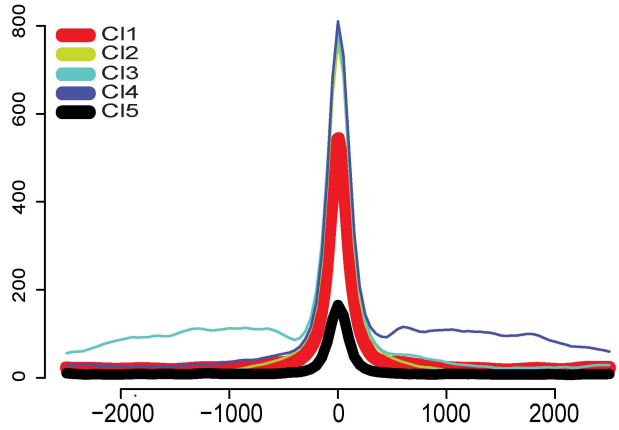
E



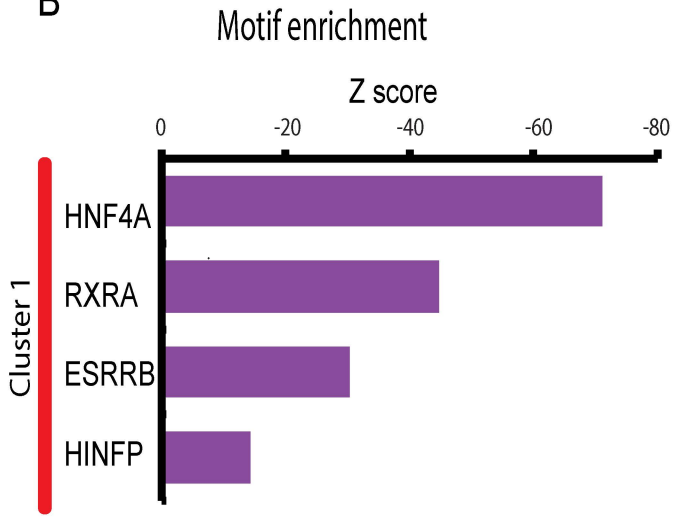
F



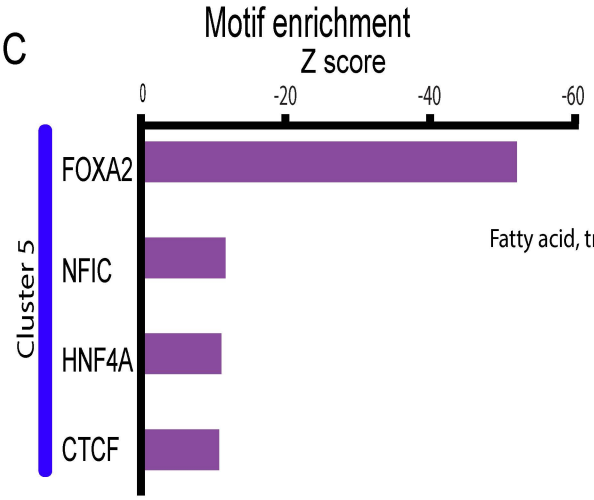
A


DNaseI profile
CI1
CI2
CI3
CI4
CI5
800
600
400
200
0
−2000
−1000
0
1000
2000

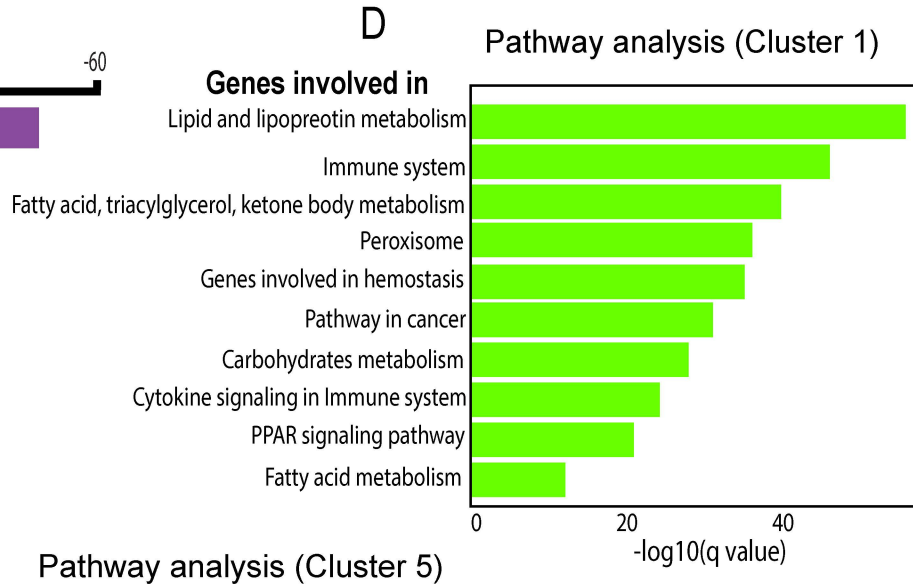
B


Motif enrichment
Z score
0
-20
-40
-60
-80
Cluster 1
HNF4A
RXRA
ESRRB
HINFP

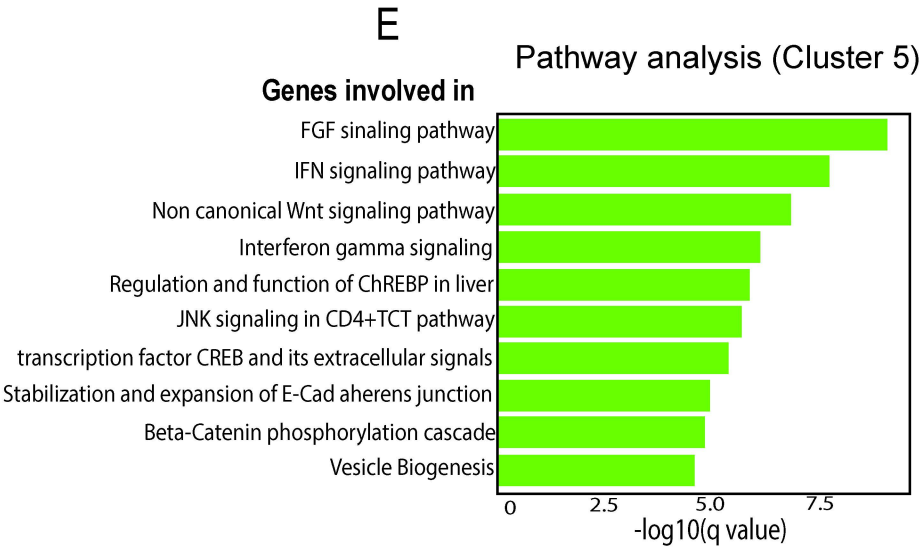
C

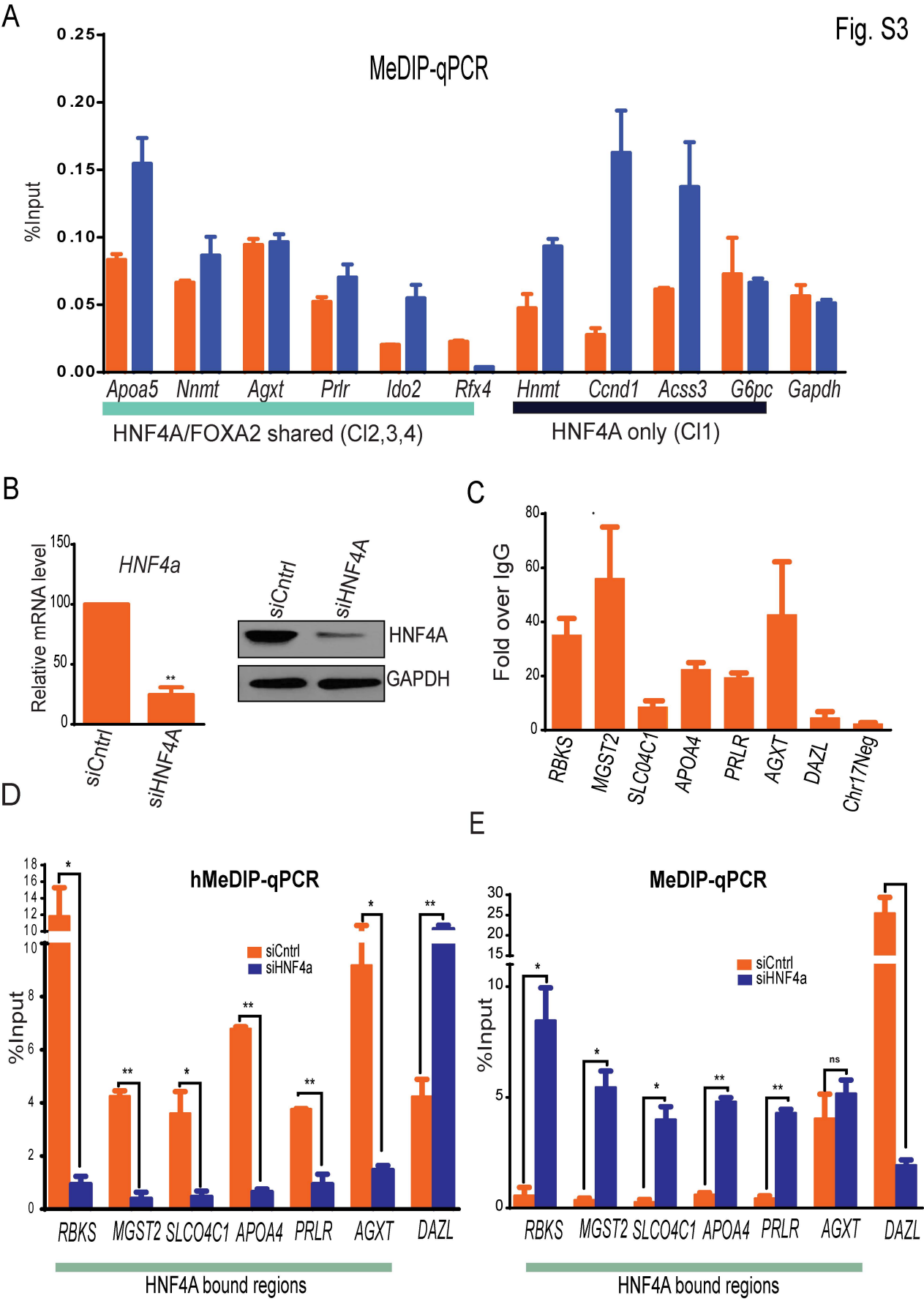

Motif enrichment
Z score
0
-20
-40
-60
Cluster 5
FOXA2
NFIC
HNF4A
CTCF

D

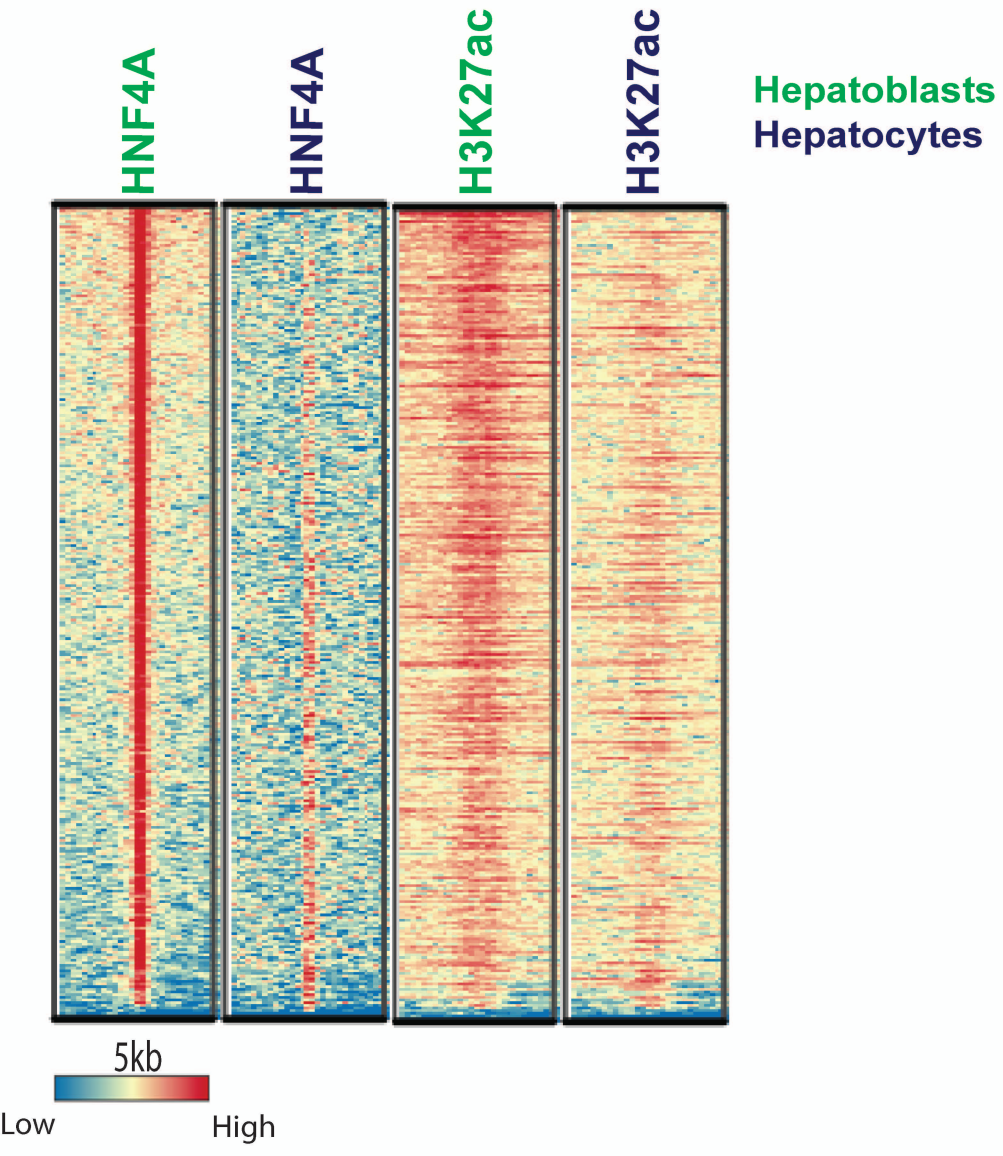

Pathway analysis (Cluster 1)
Genes involved in
Lipid and lipopreotin metabolism
Immune system
Fatty acid, triacylglycerol, ketone body metabolism
Peroxisome
Genes involved in hemostasis
Pathway in cancer
Carbohydrates metabolism
Cytokine signaling in Immune system
PPAR signaling pathway
Fatty acid metabolism
0
20
40
-log10(q value)

E

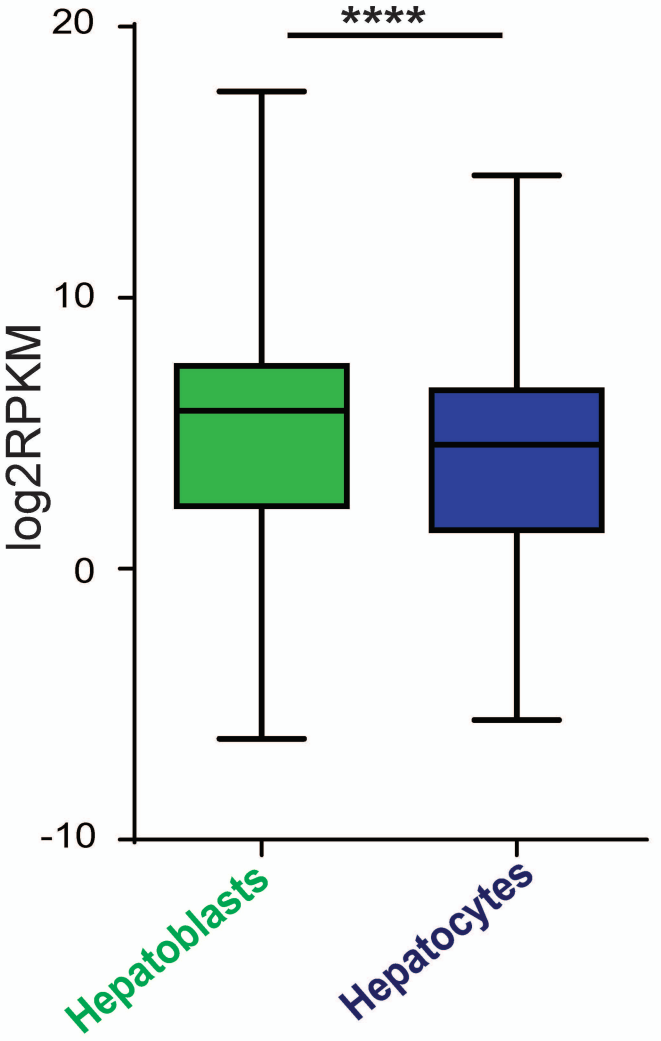

Pathway analysis (Cluster 5)
Genes involved in
FGF sinaling pathway
IFN signaling pathway
Non canonical Wnt signaling pathway
Interferon gamma signaling
Regulation and function of ChREBP in liver
JNK signaling in CD4+TCT pathway
transcription factor CREB and its extracellular signals
Stabilization and expansion of E-Cad aherens junction
Beta-Catenin phosphorylation cascade
Vesicle Biogenesis
0
2.5
5.0
7.5
-log10(q value)

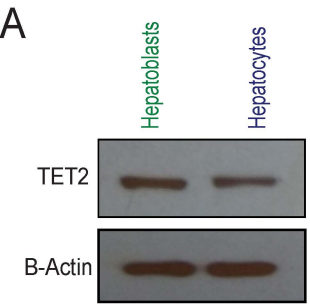

A
MeDIP-qPCR
%Input
0.25
0.20
0.15
0.10
0.05
0.00
Apoa5
Nnmt
Agxt
Prlr
Ido2
Rfx4
Hnmt
Ccnd1
Acss3
G6pc
Gapdh
HNF4A/FOXA2 shared (Cl2,3,4)
HNF4A only (Cl1)
B
Relative mRNA level
150
100
50
0
HNF4a
**
siCntrl
siHNF4A
siCntrl
siHNF4A
HNF4A
GAPDH
C
Fold over IgG
80
60
40
20
0
RBKS
MGST2
SLC04C1
APOA4
PRLR
AGXT
DAZL
Chr17Neg
D
hMeDIP-qPCR
%Input
18
16
14
12
10
8
6
4
2
0
siCntrl
siHNF4a
RBKS
MGST2
SLCO4C1
APOA4
PRLR
AGXT
DAZL
HNF4A bound regions
E
MeDIP-qPCR
%Input
30
25
20
15
10
5
0
siCntrl
siHNF4a
ns
RBKS
MGST2
SLCO4C1
APOA4
PRLR
AGXT
DAZL
HNF4A bound regions

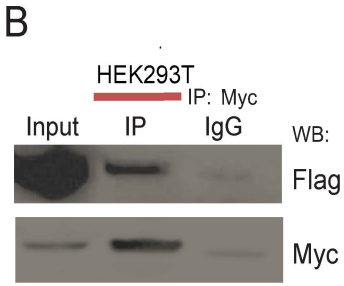
A

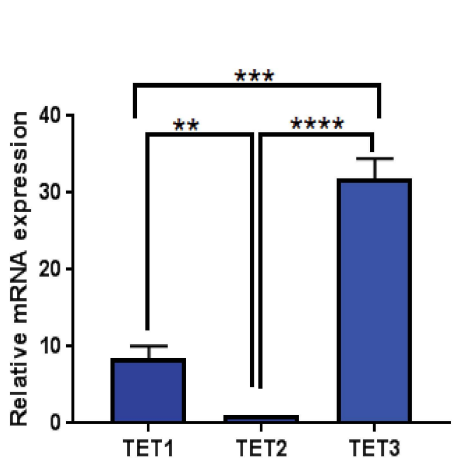

HNF4A
HNF4A
H3K27ac
H3K27ac
Hepatoblasts
Hepatocytes
5kb
Low
High

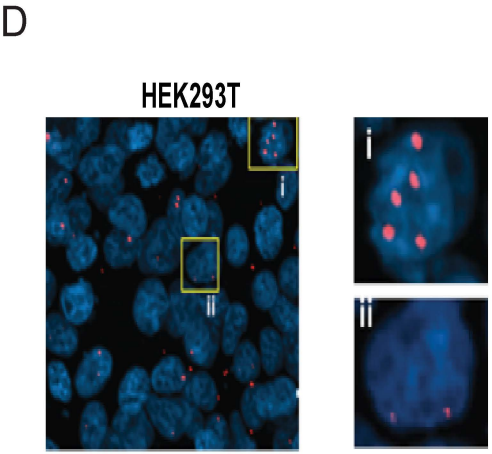
B

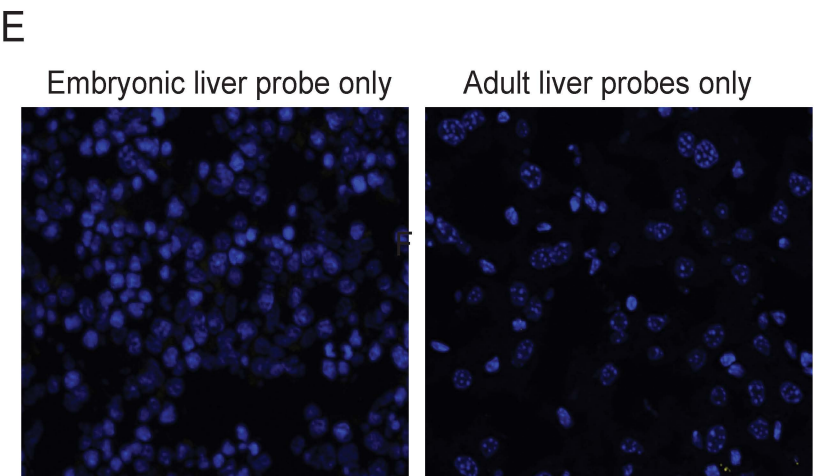

****
20
10
0
-10
log2RPKM
Hepatoblasts
Hepatocytes


A
Hepatoblasts
Hepatocytes
TET2
B-Actin

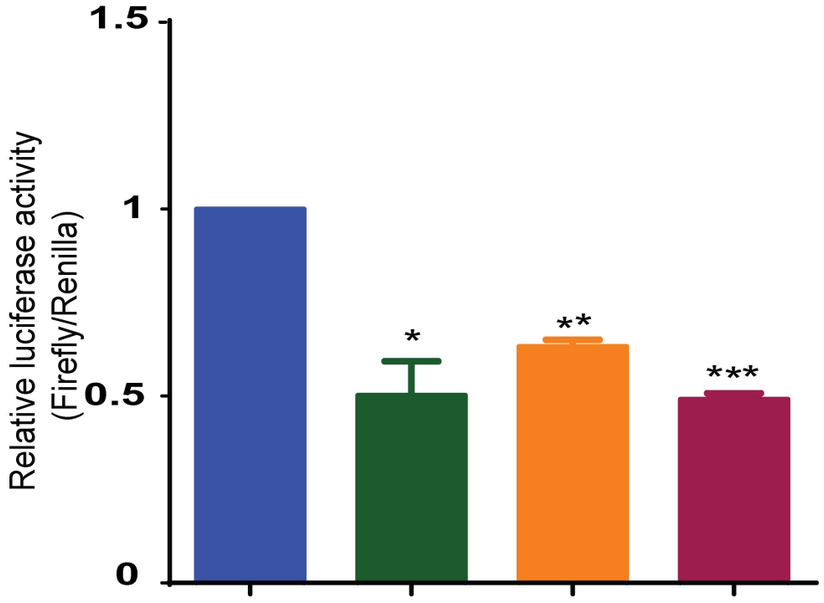

B
HEK293T
IP: Myc
Input
IP
IgG
WB:
Flag
Myc


C
***
**
****
Relative mRNA expression
40
30
20
10
0
TET1
TET2
TET3


D
HEK293T
i
ii

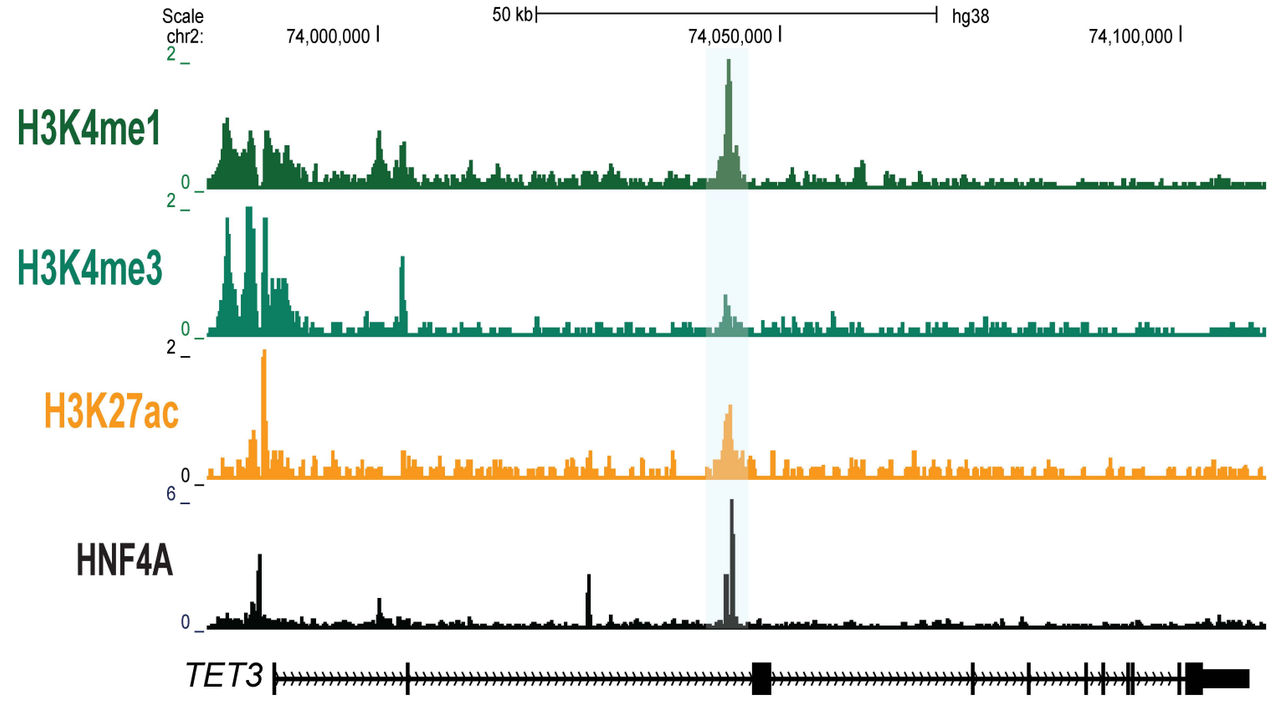

E
Embryonic liver probe only
Adult liver probes only

A

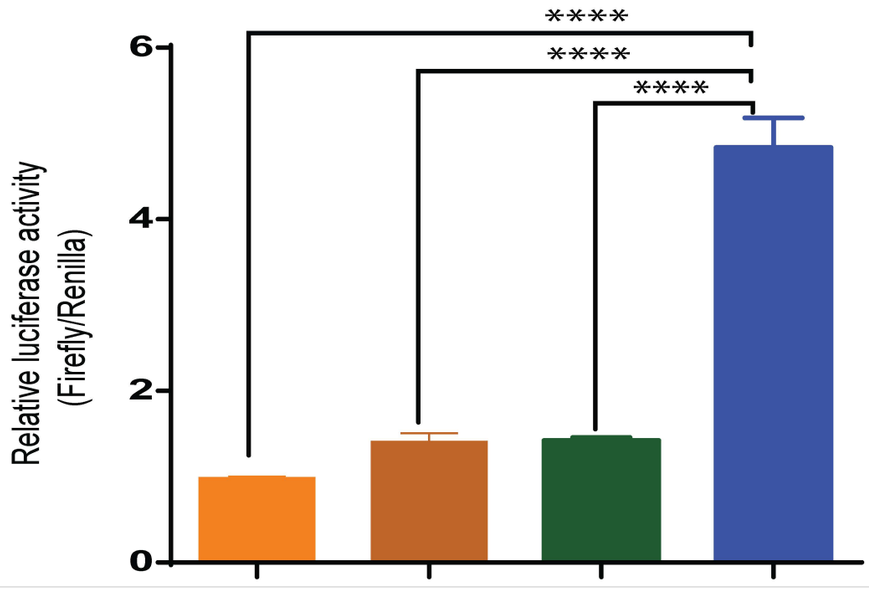


PGL3EIBP-mTET3	+	-	-	-
PGL3-EIBP-mTET3Δ1	-	+	-	-
PGL3-EIBP-mTET3Δ2	-	-	+	-
PGL3-EIBP-mTET3Δ1Δ2	-	-	-	+
PDGT-HNF4A	+	+	+	+

B



C



PGL3-EIBP-Empty	+	-	+	-
PGL3-EIBP-hTET3E	-	+	-	+
PDGT-Empty	+	+	-	-
PDGT-HNF4A	-	-	+	+

D